The Econometrics of Piecewise Linear Budget Constraints With Skewed Error Distributions: An Application To Housing Demand In The Presence Of Capital Gains Taxation

Zheng Yan

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Economics

Stuart Rosenthal, Chair
Richard Ashley
Russell Murphy
Djavad Salehi-Isfahani
Aris Spanos
Keying Ye

Aug 6, 1999
Blacksburg, Virginia

Key words: Housing Demand, Piecewise-Linear Budget Constraint, Thin Market, Skewness, Convex Budget Set, Kink, Measurement Error, Heterogeneity Error

The Econometrics of Piecewise Linear Budget Constraints With Skewed Error Distributions: An Application To Housing Demand In The Presence Of Capital Gains Taxation

Zheng Yan

(Abstract)

This paper examines the extent to which thin markets in conjunction with tax induced kinks in the budget constraint cause consumer demand to be skewed. To illustrate the principles I focus on the demand for owner-occupied housing. Housing units are indivisible and heterogeneous while tastes for housing are at least partly idiosyncratic, causing housing markets to be thin. In addition, prior to 1998, capital gains tax provisions introduced a sharp kink in the budget constraint of existing owner-occupiers in search of a new home: previous homeowners under age 55 paid no capital gains tax if they bought up, but were subject to capital gains tax if they bought down.

I first characterize the economic conditions under which households err on the up or down side when choosing a home in the presence of a thin market and a kinked budget constraint. I then specify an empirical model that takes such effects into account. Results based on Monte Carlo experiments indicate that failing to allow for skewness in the demand for housing leads to biased estimates of the elasticities of demand when such skewness is actually present. In addition, estimates based on American Housing Survey data suggest that such bias is substantial: controlling for skewness reduces the price elasticity of demand among previous owner-occupiers from 1.6 to 0.3. Moreover, 58% of previous homeowners err on the up while only 42% err on the down side. Thus, housing demand is skewed.

Acknowledgments

My most sincere thanks go to my committee chair, Stuart Rosenthal, for his inspiration, encouragement, and guidance. Without his constant and hearty attention, this dissertation would not exist.

I would also like to thank Aris Spanos, Catherine Eckel, and Richard Ashley for their helpful suggestions to this research. Helpful comments from Economic Empirical Workshop are also appreciated. I also thank Djavad Salehi-Isfahani, Russell Murphy, and Keying Ye for their support as members of my committee.

Finally, I dedicate this dissertation to my wife Rong Xiao and my parents who love me most in the world.

Table of Contents

Abstract ii
Acknowledgements iii
Table of Contents iv

Introduction 1

Chapter 1. Piecewise-linear Budget Constraint Model 5
1.1. Literature Review for Piecewise-Linear Budget Constraint Model 5
1.2. Econometric Model 7
1.3. Estimation Procedure 9
1.4. Capital Gains and Budget Constraint 12

Chapter 2. Thin Markets and Error Skewness 14
2.1. The Literature Review for Thin Markets 14
2.2. Error Skewness and Thin Markets 16
2.2.1. Linear Budget constraint 16
2.2.2. Piecewise-Linear Budget Constraint 17

Chapter 3. Monte Carlo Experiments with Simulated Data 20
3.1. Data Simulation 20
3.2. Monte Carlo Results 21

Chapter 4. Estimation with Actual Data 24
4.1. Data and Variable Description 24
4.2. Estimation Results 25

Conclusion 28

References 29

Appendix

Appendix A. Mathematical approximation to compute integral 32

Appendix B. SAS program for Monte Carlo study 34

Appendix C. SAS program for two-error normal model 45

Appendix D. SAS program for two-error normal & gamma model 49

Table

Table 1-9 54

Table 10-12 59

Table 13-15 61

Graph 1-4 66

Vita 67

Introduction

Piecewise-linear budget constraints arise in a variety of circumstances that can be classified into two major sectors: public and private. The public sector usually involves government tax and transfer programs, and the private sector involves block or discount pricing. In the public sector many empirical studies have been done in public finance and labor economics.[1] For example, the progressive income tax and a variety of welfare programs create a set of discretely varying tax rates that affect not only labor supply (Flood and MaCurdy (1992); Aronsson and Wikstrom (1994)) and after-tax wage rates (Aronsson, Wikstrom and Brannlund (1997)), but also the demand for goods such as housing (Hoyt and Rosenthal (1990), (1992)) and medical care. Another case is the program of federal grants-in-aid to states and localities. Such grants are frequently provided at discretely varying subsidy rates based on expenditure or income of the state and local governments (Barnett, Levaggi and Smith (1992)). Occasionally piecewise-linear constraints arise in the private sector when the price charged per unit depends on the bulk quantity bought by each customer. One example is the block pricing schedules in electricity and water services (Billings and Agthe (1980); Taylor (1975); Taylor, Blattenbergre, and Rennhack (1982); Terza and Welch (1982)). Another example arises in financial markets when the interest rate depends on the size of the loan or deposit.

Since piecewise-linear budget constraints arise in so many circumstances, estimation of demand (and supply) curves subject to kink budget constraints is a very important issue in the area of applied econometrics. As noted by Moffitt (1986), beginning with the study by Burtless and Hausman (1978) and later work by Hausman (1985) and others, a very general Maximum Likelihood (ML) approach was proposed to address the estimation of demand functions subject to piecewise-linear budget constraints. Moffit (1986) formalized the economic and econometric principles associated with the impact of kinked budget constraints on consumer demand. As part of that effort, Moffit derived the consumer demand function under a piecewise-linear budget constraint and provided a general ML estimation method to a two-error consumer demand model: the first error term captures heterogeneity of preferences, while the second error term controls for measurement error in the data. Moffit (1986) also derived the likelihood function

[1] See Moffitt (1986) for a survey of empirical studies prior to 1986.

under the assumption that the heterogeneity and measurement error terms are independent and distributed bivariate normal, although he suggested that the normality assumption for the error terms might be a poor approximation in finite samples. Nevertheless, nearly all applications of the kinked budget constraint model prior to and since Moffit's 1986 paper assume a bivariate normal distribution for the two error terms.[2] As will become apparent in this paper, in the context of housing demand, the assumption of bivariate normality fails to capture essential features of the housing market, especially with respect to the thinness of the market, and leads to substantially biased estimates.

A thin market is a market with few negotiated transactions per time period. Because both demand and supply in a thin market are small, prices and quantities traded may not fully reflect underlying production costs and consumer preferences. Instead, the utility maximizing quantity may not be available forcing the consumer to err on the up or down side relative to the first-best selection. I show that if the consumer is more likely to err up (or down), consumer demand will be skewed. Although such skewness does not create any problems when demand is estimated subject to a linear budget constraint, that does not carry over to the case of kinked budget constraint models. When the budget constraint is kinked, skewed measurement errors cause ML estimation based on normally distributed error terms to yield biased and inconsistent estimates.

In the owner-occupied housing market prior to 1998, previous homeowners faced a piecewise-linear budget constraint because of capital gains tax provisions: previous owner-occupiers under age 55 did not have to pay capital gains tax if they bought a home of equal or greater value to their previous home, but did pay capital gains tax if they bought down. Because housing units are indivisible and heterogeneous and tastes for housing are at lest partly idiosyncratic, housing markets are thin. I show that when the household's preferred housing bundle is not available, the household is more likely to err on the up side as the tax penalty from buying down increases (sharpening the kink).[3] As a consequence, the measurement error term in the two-error housing demand model is likely to be skewed to the left in contrast to traditional

[2] Hausman (1985) tested both a truncated normal and a Weibull distribution for his heterogeneity error in a two-error model, and found little difference in the coefficient estimates.

[3] In addition, Hoyt and Rosenthal (1990) suggest that realtors working on a commission have an incentive to direct the consumer towards properties that sell at or above the value of the previous home. This would further skew the demand for housing among families seeking to purchase a new home of equal value to the old home.

kinked budget constraint models that assume normality.[4]

To address these issues, I first characterize the economic conditions that govern whether a household will err on the up or down side when the preferred (first-best) selection is not available. I then conduct a series of Monte Carlo experiments for a two-error housing demand model and test the sensitivity of results to skewness in the measurement error term. The performance of both ML and Ordinary Least Square (OLS) estimators are examined under a variety of sample size, error variance, error skewness, and kink angle. Results from the Monte Carlo experiments indicate that: (1) both sample size and error variance have a significant effect on the dispersion of the estimated elasticities of demand for housing; (2) ML estimates are unbiased, consistent, and superior to OLS estimates when there is a kinked budget constraint and the measurement error term is normally distributed; (3) ML estimates based on the assumption of normally distributed error terms are biased and inconsistent when the measurement error term is truly skewed.

I also estimate housing demand using American Housing Survey (AHS) data based on the two-error specification allowing for skewness and a kinked budget constraint. Findings from that model are compared to estimation results from a traditional kinked budget constraint model that assumes bivariate normality. Results indicate that failing to allow for skewness in the measurement error term leads to substantially biased estimates of the elasticities of demand for housing: controlling for skewness reduces the price elasticity of demand among previous owner-occupiers from 1.6 to 0.3. In addition, 58 percent of previous homeowners err on the up side while 42 percent err on the down side when buying a new home. Thus, housing demand is skewed.

The remainder of the paper is organized as follows. In chapter 1, the piecewise-linear budget constraint model and the estimation procedure are presented. Chapter 2 provides the theoretical motivation for error skewness in the presence of kinked budget constraints and thin markets. In chapter 3, I analyze the results of Monte Carlo experiments under varying sample size, error variance, error skewness, and kink angle, which test the sensitivity of results to error skewness. In chapter 4, I estimate housing demand allowing for a kinked budget constraint and skewed measurement error term and compare results to those based on a model in which the

[4] This intuition came from Hoyt and Rosenthal (1990). They defined a skewed kink range and used a heterogeneity-error-only model to estimate housing demand.

heterogeneity and measurement error terms are assumed to be distributed bivariate normal. A conclusion section follows.

Chapter 1. Piecewise-Linear Budget Constraint Model

1.1. Literature Review for Piecewise-Linear Budget Constraint Model

Some simple statistical methods were used to estimate the piecewise-linear budget constraint problem before 1980, but their performance was questioned in the middle of 1980's (see Moffitt (1986)). In the case of the linear budget constraint, simple estimation methods (e.g., least-squares and instrumental-variable procedures) give consistent and unbiased estimators because the consumer is assumed to purchase any desired quantity at a constant price subject to a budget constraint. However, in the case of piecewise-linear budget constraint, because the price paid at the margin depends on the quantity of the goods consumed and random components (heterogeneity error and measurement error) exist in the determination of the quantity, there will be nonzero covariance between the error term and the variables measuring the marginal price and effective or virtual income. Applying Ordinary Least Squares (OLS) produces biased and inconsistent estimators of the parameters of interest. Note that the two-step estimation method for sample selection models is not applicable to the two-error piecewise-linear budget constraint model because the two-step method requires sample separation in the likelihood function. However, if the heterogeneity-error-only model is assumed, which implies that the desired segment and kink location are known, the likelihood function separates, and the two-step method can be applied. As noted by Moffitt (1986), while the measurement-error-only model was criticized for requiring that there is only one utility-maximizing choice in the population, the heterogeneity-error-only model was also criticized for requiring the econometrician know the exact location of the kink value in the budget constraint of each sample member.

Beginning with the study by Burtless and Hausman (1978) and some later work by Hausman (1985) and other economists as noted by Moffitt (1986), a very general econometric solution, Maximum Likelihood (ML) estimation, was demonstrated to the piecewise-linear budget constraint problem. But the formal econometric model and econometric methods in a simpler and more general fashion had not been set out until Moffitt (1986). In his article, Robert Moffitt formally derived the consumer demand function under piecewise-linear constraint and provided us the general ML estimation method to a two-error consumer demand model. Moffitt's article serves as the basic reference to the future users of these techniques. MaCurdy, Green, and Paarsch (1989) challenged the ML estimation of consumer choice problem with nonlinear budget sets. They demonstrated that ML estimation implicitly relies on the satisfaction of inequality constraints on parameters and that these constraints arise, not as a consequence of economic theory, but instead as a requirement to create a properly defined statistical model. Although these implicit constraints play a major role in explaining the discrepancies in estimates found in the literature on workers' labor supply, there is no general belief about whether the parameter restriction problem could apply to the empirical analysis outside labor economics.[5]

As noted by Moffitt (1986), the sensitivity of the ML Estimates to various forms of specification error, especially in finite samples, is an important issue in the estimation of the piecewise-linear-constraint model. Moffitt (1986) derived the likelihood function when the heterogeneity and measurement errors are both assumed to be normally distributed, although he also suggested that the normality assumption for the error terms might be a poor approximation in finite samples. Nevertheless, nearly all applications of the kinked budget constraint model prior to and since Moffit's 1986 paper assume a bivariate normal distribution for the two error terms. But in some circumstances, for example in the presence of thin markets, bivariate normally distributed error terms will not be a reasonable assumption in the two-error demand model. Since a thin market operates with few negotiated transactions per time period, both the prices and quantities of trade may not fully reflect underlying production costs and consumer preferences. Instead, the utility maximizing quantity may not be available forcing the consumer to err on the up side or on the down side relative to the first-best selection. I show that if the consumer is more likely to err up (or down), consumer demand will be skewed. Although the skewed measurement error term does not cause any problem in the estimation of linear budget constraint model, it does in the piecewise-linear budget constraint model. When the error term is truly skewed, ML estimation based on the normally distributed error term will lead to biased and inconsistent estimates.

Although there is much empirical and theoretical analysis about estimation of consumer demand under piecewise-linear budget constraint, evaluation of these estimators based on Monte Carlo experiments has been limited. Sharon Megdal (1987) examined the performance of two ML estimators and one OLS estimator under varying sample sizes and error variances to a heterogeneity-error-only consumer demand model. Robert Triest (1987) investigated the

[5] Generally it is reasonable to assume that the utility function is strictly quasi-concave outside labor economics. But in labor economics it isn't because of the possibility of a backward binding labor supply curve.

performance of the ML estimator and two simple estimators (OLS and instrumental-variable procedures) when there was measurement error in the budget constraints of sample members. In chapter 3, I present the results for a series of Monte Carlo experiments for a two-error demand model when the budget constraint is piecewise-linear and the budget set is convex.

1.2. Econometric Model

Suppose that a consumer maximizes a strictly quasi-concave utility function U (X, Y; α, β) subject to a piecewise-linear budget constraint, where X and Y are two goods, α and β are the preference parameters, and the price of Y is used as a numeraire. The demand function for X resulting from utility maximization can be written in the general form X = g (P, M; α, β), where P and M are the price of X and consumer's income respectively. Consider a piecewise-linear convex budget set with two segments, as shown in Figure 1. The budget constraint can be written as

$$M_1 = P_1X + Y, \qquad \text{if } X \leq X^*,$$
$$M_2 = P_2X + Y, \qquad \text{if } X > X^*. \tag{1}$$

where X^* is the kink value, P_i is the price on segment i and $P_1 < P_2$, M_2 is the virtual income, and $M_2 = M_1 + (P_2 - P_1) X^*$. Given this budget constraint, if an interior optimum exists, it is unique and occurs at either a tangency with a segment or at the kink value (see Hausman (1985) and Moffitt (1986)). In this case, the consumer demand function is

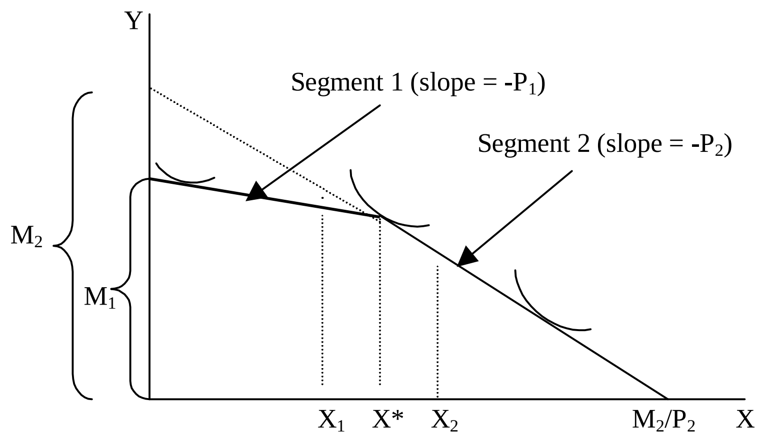


Figure 1. Piecewise-Linear Convex Budget Set

$$X = g(P_1, M_1; \alpha, \beta), \quad \text{if } g(P_1, M_1; \alpha, \beta) < X^*, \quad \text{(Segment 1)}$$
$$= X^*, \quad \text{if } g(P_2, M_2; \alpha, \beta) \leq X^* \leq g(P_1, M_1; \alpha, \beta), \quad \text{(Kink)}$$
$$= g(P_2, M_2; \alpha, \beta), \quad \text{if } X^* < g(P_2, M_2; \alpha, \beta), \quad \text{(Segment 2)} \qquad (2)$$

In empirical analysis, consumer demand is often specified as a log-linear function of price and income. Socio-demographic variables are also typically included in the empirical model as taste shifters, while α represent the randomness in preferences that are not captured by inclusion of those variables. This yields the following specification for consumer demand:

$$\log X_i = a_0 + a_1 \log P_i + a_2 \log M_i + \textstyle\sum_j Z_{ij}\, \delta_{ij} + \alpha_i + \varepsilon_i, \qquad (3)$$

where α_i and ε_i are heterogeneity error and measurement error respectively, P_i and M_i are assigned values corresponding to the household's preferred (first-best) budget segment, and Z_i is a vector of household socio-demographic characteristics. Simplifying notation, the demand function becomes:

$$x_i = z_{1i} + \alpha_i + \varepsilon_i, \quad \text{if } z_{1i} + \alpha_i < x_i^*, \quad \text{(Segment 1)}$$
$$= x_i^* + \varepsilon_i, \quad \text{if } z_{2i} + \alpha_i \leq x_i^* \leq z_{1i} + \alpha_i, \quad \text{(Kink)}$$
$$= z_{2i} + \alpha_i + \varepsilon_i, \quad \text{if } z_{2i} + \alpha_i > x_i^*, \quad \text{(Segment 2)} \qquad (4)$$

where $x = \log X_i$, $x_i^* = \log X_i^*$, $z_{1i} = a_0 + a_1 \log P_{1i} + a_2 \log M_{1i} + \sum_j Z_{ij}\, \delta_{ij}$, and $z_2 = a_0 + a_1 \log P_{2i} + a_2 \log M_{2i} + \sum_j Z_{ij}\, \delta_{ij}$. Note that because of measurement error, the segment on which an observation is observed is not necessarily the preferred (first-best) budget segment. Instead, either because of coding error in the data or thin market effects discussed earlier, it is possible that the observed choice of budget segment differs from the household's true preferred choice.

Before proceeding further, it is useful to highlight several points about α and ε. First, because the values of P and M depend on the preferred (first-best) value of X in (3), the error term α is positively correlated with the price and income variables. As is well established that correlation causes OLS estimates of a_0, a_1, and a_2 to be biased and inconsistent if we ignore the kinked budget constraint.

Second, heterogeneity of preferences generates clusters of observations at the kink point of a convex budget constraint (see Moffitt (1986)). The economic interpretation of the clustering is straightforward: from (4) we know that $x^* + z_1 \leq \alpha \leq x^* + z_2$ at the kink. Thus, there will be an entire range of utility functions compatible with utility maximization at x^* (see Figure 5 in Moffitt (1986)).

Third, and most important for this paper, measurement error ε includes not only coding error, but also optimization error [e.g. Hausman (1985)].[6] Whereas coding error is likely to be symmetrically distributed, for reasons outlined in the Introduction, optimization error is likely to be skewed. Moreover, as will become apparent, the sharper the kink, the more skewed the distribution of ε.

1.3. Estimation Procedure

In empirical studies, the error terms α and ε are generally assumed to be independent and follow a bivariate normal distribution with mean zero and variances σ_α^2 and σ_ε^2 respectively. Knowing the distribution of the error terms, we could estimate the demand function using ML methods, which have desirable large sample properties but unknown small sample properties. The likelihood function contains the probabilities for each X value, and each probability is the sum of the joint probabilities that an individual is maximizing utility on each segment or kink:

$$L = \Pi_{i=1}^{n} \Pr (x_i) \tag{5}$$

where

$$\begin{aligned} \Pr (x_i) = &\Pr [v_i = x_i - z_{1i}, \alpha_i < x_i^* - z_{1i}] && \text{(Segment 1)} \\ &+ \Pr [v_i = x_i - z_{2i}, \alpha_i > x_i^* - z_{2i}] && \text{(Segment 2)} \\ &+ \Pr [\varepsilon_i = x_i - x_i^*, x_i^* - z_{1i} \leq \alpha_i \leq x_i^* - z_{2i}] && \text{(Kink)} \end{aligned} \tag{6}$$

where $v_i = \alpha_i + \varepsilon_i$.

Note that there is no sample separation in the likelihood function; because of measurement error, true segment and kink choices are unobserved. This is also why the two-step estimation methods for sample selection models are not applicable to the two-error piecewise-linear budget constraint model, because the two-step method requires sample separation in the likelihood function. However, if the heterogeneity-error-only model is assumed, which implies that the desired segment and kink location are known, the likelihood function separates, and the two-step method can be applied.

In the context of housing demand, however, the estimated price and income elasticity of demand from (5) and (6) might be sensitive to the distribution imposed for ε.[7] For example,

---

[6] In contrast, Moffitt (1986) characterizes both sources of error as measurement error.

[7] Hoyt and Rosenthal (1990) made this point as well but did not formally incorporate non-normal error terms into

suppose a consumer maximizes utility near the kink at segment 1. If the distribution of ε is skewed to the left, when evaluating the family's choice of X, more weight will be placed on the price and income terms associated with segment 2, and less weight will be given to segment 1, at least as compared to models based on a symmetric distribution for ε. Since skewness affects the weight placed on the different price and income terms, ML estimates based on (symmetric) normal error terms will be biased and inconsistent when the data are actually skewed.

The degree to which the budget constraint is kinked -- referred to as the kink angle for the remainder of the paper -- also affects OLS and ML estimates of the demand function. For my purpose, the angle of the kink is measured by $P_2 / P_1 - 1$, where P_1 and P_2 are the slopes of the budget segments in Figure 1. When the kink angle = 0 or the budget constraint is linear, ML and OLS estimates are equivalent and both are BLUE (Best Linear Unbiased Estimator). Assuming ε is normally distributed, ML estimates are unbiased, and work better as opposed to the OLS estimates as the kink angle increases, because ML estimation recognizes the kink. But when the error term is truly skewed, ML estimates, based on the assumption of normal error terms, are biased and inconsistent, and the larger the kink angle, the more biased and inconsistent the ML estimates.

An important contribution of this paper is to explicitly allow for skewed distributions for ε when estimating the kinked budget constraint model. Moffitt (1986) derived the likelihood function assuming α and ε are independent and follow a bivariate normal distribution. Under these assumptions, as shown in the Appendix to Moffit's (1986) paper, the bivariate distribution of v and ε can be transformed into the product of two separate independent normal distributions. But when we allow ε following a skewed distribution such as the Gamma distribution, the marginal density function of v does not have a closed form solution, and the bivariate density of v and α can not be separated into two independent distributions. To address that problem, I developed a numerical approximation for the integral of the bivariate density (see Appendix 1). In order to increase the speed of convergence, equation (5) was maximized using the Nelder-Mead simplex algorithm.[8]

---

the likelihood function.

[8] The Nelder-Mead algorithm changes the shape of the simplex by adapting the non-linearity of the objective function, which contributes to an increased speed of convergence.

For the housing application to follow, I argue ε is skewed to the left to the extent that tax considerations encourage previous house owners to err on the up side. For that reason, when specifying the likelihood function, I impose the assumption that ε is drawn from a left skewed distribution. To operate the model, I therefore assume that ε follows the symmetric opposite of the Gamma distribution -- referred to as the Reversed Gamma distribution hereafter.[9] The Gamma distribution is well behaved and right skewed. Reversing the Gamma distribution allows us to estimate the model, while also imposing left skewness on ε.[10] The density function for ε is $f(\varepsilon) = ab - \varepsilon^{a-1} e^{-\varepsilon/b} / b^a \Gamma(a)$ for $\varepsilon < ab$ and f (ε) is skewed to the left with mean 0, variance ab^2, and skewness -2/sqrt(a).[11] When a is large, ε is approximately normal with mean 0 and variance ab^2.[12] Suppose α still follows a normal distribution with mean 0 and variance σ^2, then the bivariate distribution of v and α can be expressed as:

$$f_{v,\alpha}(v, \alpha) = f_{u,\alpha}(ab + \alpha - v, \alpha) = g(ab + \alpha - v)\,\phi(\alpha/\sigma)/\sigma \qquad (7)$$

where (ab + α - v) > 0 or α > (v – ab), g is the Gamma density function and ϕ is the standard normal density function. Then

$$\Pr(X_i) = \text{Prob}_{1i} + \text{Prob}_{2i} + \text{Prob}_{3i}$$

$$\text{Prob}_{1i} = \int_{v1-ab}^{\alpha 1} [g(ab + \alpha - v1)\,\phi(\alpha/\sigma)/\sigma]\,d\alpha \qquad \text{(buy down)}$$

$$\text{Prob}_{2i} = \int_{\alpha 2}^{\infty} [g(ab + \alpha - v2)\,\phi(\alpha/\sigma)/\sigma]\,d\alpha \qquad \text{(buy up)}$$

$$\text{Prob}_{3i} = g(ab - s) \int_{\alpha 1}^{\alpha 2} [\phi(\alpha/\sigma)/\sigma]\,d\alpha \qquad \text{(locate on kink)} \qquad (8)$$

where $v1 = X_i - X_{1i}$, $v2 = X_i - X_{2i}$, $\alpha 1 = X_i^* - X_{1i}$, $\alpha 2 = X_i^* - X_{2i}$, and $s = X_i - X_i^*$.

---

[9] Let u = ab - ε, then u follows a Gamma distribution with density function $g(u) = u^{a-1} e^{-x/b} / b^a \Gamma(a)$ for u > 0. g (u) is skewed to the right with mean ab, variance ab^2, and skewness 2/sqrt(a).

[10] Although Beta distribution allows symmetry and skewness to both left and right, assuming Beta distributed ε does not lead to an operational model because ML estimation hardly converge.

[11] The skewness of ε is defined as $E(\varepsilon^3)/[\text{var}(\varepsilon)]^{3/2}$.

[12] Here I give a simple proof for Gamma distribution. It is not difficult to show that the same conclusion applies to Reversed Gamma distribution since Reversed Gamma distribution is just a linear transformation of Gamma distribution. Suppose a > n and u_i follow a gamma distribution with parameters a/n and b or u_i ~ gamma (a/n, b) for i=1,2,…,n, and all u_i's are independent from each other. From the characteristics of Gamma distribution we know $\Sigma_{i=1}^n u_i$ ~ gamma (a,b). Also according to the Central Limit Theory we know that if $\Sigma_{i=1}^n u_i / n$ is approximately normal $(ab/n, ab^2/n^2)$ when n is large, then $\Sigma_{i=1}^n u_i$ is approximately normal (ab, ab^2) when n is large. So gamma (a, b) is approximately normal (ab, ab^2) when a is large (recall a > n).

Among $Prob_{1i}$, $Prob_{2i}$, and $Prob_{3i}$, $Prob_3$ is easy to compute since the bivariate density of v and ε has been separated to two independent densities of α and ε. But for the other two, we need some mathematics approximation to compute the integral (see Appendix 1).

1.4. Capital Gains and Budget Constraint

Prior to 1998, capital gains tax provisions introduced a sharp kink in the budget constraint of existing owner-occupiers in search of a new home.[13] The capital gain tax for previous house owner who is under 55 and buy down is determined by

$$T = \min [k*t*G, k*t (PVAL - H*P_h)],$$

where t is the marginal income tax rate, k capital gains tax rate, G the capital gain, PVAL the sale price of previous home, and $H*P_h$ the current house value.

Annualizing the capital gains tax over the household's lifetime (by multiplying T by the real interest rate r), the family's annual budget constraint can be written as

$$\begin{aligned} Y &= X + H*R_0, && H*P_h \geq PVAL, \\ Y - k*t*r (PVAL - H*P_h) &= X + H*R_0, && PVAL - G < H*P_h < PVAL, \\ Y - k*t*r*G &= X + H*R_0, && H*P_h < PVAL - G, \end{aligned} \quad (9)$$

where the rental cost of housing services (R_0) is defined as the cost to an owner-occupier of one dollar of housing services in the rental market.

Note that the taxation for capital gains actually creates two kinks in the budget constraint, one convex kink and one concave kink, as shown in figure 2. In my data, I have information on all of the parameters of the budget constraint except G, the capital gain on the previous home. For this reason, throughout my analysis I assume that families do not locate below (PVAL - G)/P_h, and instead, I focus my discussion on the distortion induced by the convex kink at PVAL/P_h.

According to the tax code, kink value is defined in terms of housing expenditure and equal to PVAL. This is somewhat unusual and differs from the model specified in section 1.2 because X* is the quantity of the goods or hours in the labor supply model. But if we assume that

[13] Homeowners do not pay capital gains tax when they move from their home if they purchase another home of equal or greater value within two years. After the age of 55, homeowners are also allowed a one-time exemption from payment of capital gains tax for the first $125,000 of capital gains. But when homeowners are under 55 years old and buy down, they do pay capital gains tax. Prior to 1986, capital gains were taxed at 40% of the marginal income tax rate; but following TRA86, capital gains were taxed at 100% of the marginal income tax rate. After 1998, capital gains tax in owner occupied housing was removed.

housing stock can be divided into equal homogeneous units and each of them sell for the same price (P_h), we can have the usual kink value X* which is equal to PVAL/ P_h.

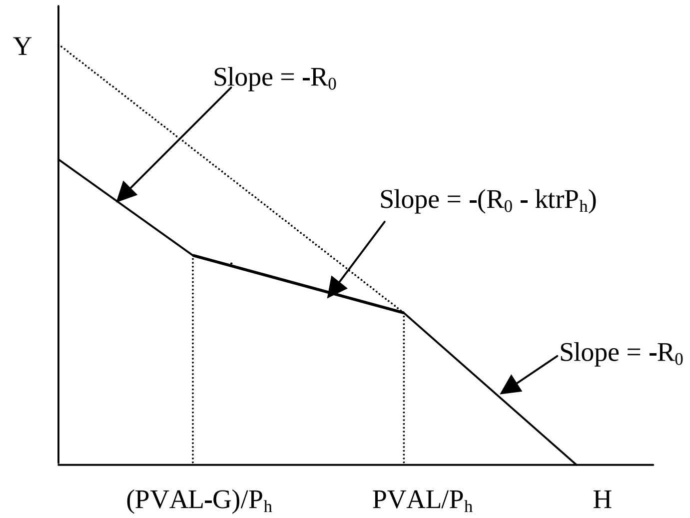


Figure 2. The Kinked Budget Set for Previous House Owner under 55

Clearly, ignoring the kinked budget constraint of previous owners will produce biased estimates of the elasticities of demand. Consider a previous homeowner who chooses to locate at the kink and recall that the kink is defined based on the nominal value of the previous home. Then for a small change in the unit price of housing, expenditure on X remains unchanged, and the quantity of X adjusts in a proportional and opposite direction. As a result, the price elasticity of demand for housing is biased towards -1. Similarly, the income elasticity is biased to zero because expenditure on X is insensitive to small changes in income for the consumer at the kink (e.g. Hoyt and Rosenthal (1990)).

Chapter 2. Thin Markets and Error Skewness

2.1. Literature Review for Thin Markets

Since Gray (1960) popularized the term "thin market" in the context of futures markets, "thin market" has been widely used in finance (Glaeser and Kallal (1997); Martikainen and Puttonen (1994); Pagano (1989)), agriculture economics (Nelson and Turner (1995); Hayenga 1978)), housing economics (Anas (1997); Arnott (1988); Sass (1987)), labor economics (Pissarides (1992)), and industry organization & Game Theory (Saleth, Braden, and Eheart (1991); Caves (1978)). Although many markets have been classified as "thin markets", for example financial markets, housing markets, various kinds of agricultural and foods markets (see Hayenga (1978)), resource markets, infant markets, terminal markets, and auction markets, there is not a formal and widely accepted definition for this term. Let us first take a look at some representative discussions about "thin market". This might help to clarify this ambiguous concept.

At a conference entitled the "Symposium on Pricing Problems in the Food Industry (with Emphasis on Thin Market)" held in Washington D.C., Hayenga, Gardner, Paul, and Houck (1978) gave a definition for thin markets. Thin markets are "markets with little trading volume and liquidity in which individual firms or offers to buy or sell can sometimes exert 'undue' influence on price or other terms of trade, usually to the detriment of others, but sometimes themselves". (p. 7)

Also in their paper, Hayenga, Gardner, Paul, and Houck (1978) gave another definition for thin markets. Although "there is a general agreement on the need for language to characterize both the (1) fewness of negotiated trades in a specified market and time period, and (2) the level of market performance, especially its liquidity and corresponding price sensitivity to incremental buy or sell orders", they finally "fell that a separation of these two criteria seems appropriate. Thus, a 'thinly-traded' market, or 'thin' market for short, would be a market with few negotiated transactions per time period." (p. 11)

Tomek and Robinson (1990) described the conditions that give rise to thin markets "as the volume sold through central markets becomes smaller, the prices established on such markets may not fully reflect aggregate supply and demand conditions; furthermore, they are more susceptible to manipulation. This is commonly referred to as the 'thin market' problem. A thin market is one in which relatively few transactions establish prices."

Saleth, Braden, and Eheart (1991) defined thin market as "a market with few eligible participants. Thin markets are likely to be manipulated by participants, leading to inefficiency."

Apparently all the literatures have focused on the structure and performance of thin markets. For one thing, thin markets can be characterized by their structure properties, for instance, few participants (buyer, seller, or both), few (negotiated) transactions, or little trading volume in a time period. On the other hand, thin markets may also be characterized by their performance, which is best measured by "their possible failures to provide a reasonably stable, efficient, equitable, minimal risk means of transferring ownership of a product from buyer to seller, and provide prices which are accurate signals of supply-demand conditions, particularly due to insufficient market liquidity or possible market manipulation." (Hayenga, Gardner, Paul, and Houck 1978, p. 10)

Although I do not have strong reasons, I think defining thin markets based solely on their structure properties is appropriate.[14] Moreover, among the three factors that characterize the structure of thin markets, I think the number of transactions serves best to define this concept. Due to reference price policy widely used in contracts and "follow the leader" price policy adopted by the firm in a competitive disadvantage position, not all the transactions are actually or potentially involved in the price determination process and negotiated on an open or public market. Therefore, the number of negotiated transactions is the key criteria in defining thin market.

In summary, I agree with the second definition by Hayenga, Gardner, Paul, and Houck (1978): "a 'thinly-traded' market, or 'thin' market for short, would be a market with few negotiated transactions per time period." (p. 11)

Thin markets arise in a variety of circumstances:

- vertical integration, long term contractual commitments, or reference price contracts let the residual market become thin (Hayenga, Gardner, Paul, and Houck (1978); p. 8)
- heterogeneous goods and partly idiosyncratic preferences (Anas (1997))[15]

---

[14] Two tentative reasons are (1) ideally, a definition for a market should be based only on its structural properties, and (2) a market with few participants, few transactions, or little trading volume may not perform badly.

[15] In Monte Carlo studies and empirical analysis, I use a housing market example. Because housing units are indivisible and heterogeneous and tastes for housing are partly idiosyncratic, housing markets are thin and private

- infant or terminal market (Hayenga, Gardner, Paul, and Houck (1978); p. 8)
- institutional barriers or artificial restrictions (Hayenga, Gardner, Paul, and Houck (1978); p. 12); high costs of arbitrage over space or over time (Caves (1978); p. 20)
- scale economies or discontinuities or fixed costs in the transaction (Caves (1978); p. 20)
- poor information and/or poor communication facilities (Powers (1978); p.35)

2.2. Error Skewness and Thin Markets

A thin market is a market with few negotiated transactions per time period. Because both demand and supply in a thin market are small, prices and quantities traded may not fully reflect underlying production costs and consumer preferences. Instead, the utility maximizing quantity may not be available forcing the consumer to err on the up side or on the down side relative to the first-best selection. This section characterizes the economic conditions that govern whether a household under both linear and piecewise-linear budget constraints will err on the up or down side when the preferred selection is not available.

2.2.1. Linear Budget constraint

Consider a linear budget set as shown in Figure 3, where X_0 is the utility maximization value, P is the price of X, and M is income. Because the market is thin, X_0, the consumer's preferred choice, is not available, but two other options $X_1 = X_0 - d$ and $X_2 = X_0 + d$ are available where d is a small positive number compared to X_0. What level of X should the consumer choose?

Suppose the utility corresponding to X_1 is $U_1 = U(X_1, Y_1; \alpha, \beta)$, and the utility corresponding to X_2 is $U_2 = U(X_2, Y_2; \alpha, \beta)$, in which $Y_i = M - X_i*P$ (i=1, 2). The consumer will choose X_2 when $U_1 < U_2$. If the consumer is more likely to choose X_2 or err on the up side, then the error term in the demand function will be the left. In order to derive the economic conditions that govern whether a household will err on the up side or the down side, I assume a Cobb-Douglas utility function $U(X, Y; \alpha, \beta) = X^{\alpha}Y^{\beta}$. Solve for utility maximization we have

$$X_0 = M/[P*(1 + \beta/\alpha)] \quad (10)$$

search is costly, time consuming, risky and information intensive.

Let's take the first order Taylor expansion for U_1 and U_2 around X_0.

$$U_1 = X_1^{\alpha}Y_1^{\beta} \approx X_0^{\alpha}(M - X_0{*}P)^{\beta} - \alpha X_0^{\alpha-1}(M - X_0{*}P)^{\beta}d + \beta X_0^{\alpha}(M - X_0{*}P)^{\beta-1}P{*}d$$

$$U_2 = X_1^{\alpha}Y_1^{\beta} \approx X_0^{\alpha}(M - X_0{*}P)^{\beta} + \alpha X_0^{\alpha-1}(M - X_0{*}P)^{\beta}d - \beta X_0^{\alpha}(M - X_0{*}P)^{\beta-1}P{*}d \quad (11)$$

Then $U_1 < U_2 \Leftrightarrow X_0 < M/[P{*}(1 + \beta/\alpha)]$ (12)

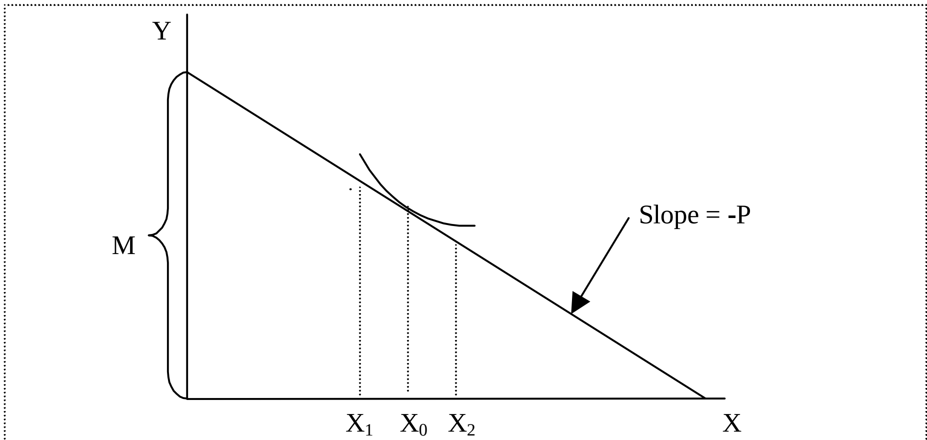


Figure 3. Linear Budget Constraint

Note that equation (12) contradicts equation (10), so consumers will be equally likely to err up or down when X_0, the first best solution, is not available. Therefore, the measurement error is symmetric. Moreover, in a linear budget constraint model, OLS will always give BLUE estimates no matter whether the error term is skewed or not.

2.2.2. Piecewise-Linear Budget Constraint

Consider a piecewise-linear convex budget set with two segments, as shown in Figure 1. The constraint can be written as:

$M_1 = P_1X + Y,$ if $X \le X_0$,

$M_2 = P_2X + Y,$ if $X > X_0$, (13)

in which X_0 represents the kink value X* where the consumer maximizes his utility, P_i is the price on segment i, $P_1 < P_2$, M_2 is the virtual income, and $M_2 = M_1 + (P_2 - P_1){*}X_0$. Because the market is thin, X_0 is not available, but two other options, $X_1 = X_0 - d$ and $X_2 = X_0 + d$, are provided, in which d is a small number comparing to X_0. Assume Y_0, Y_1, and Y_2 are the corresponding values for X_0, X_1, and X_2, then $Y_1 = Y_0 + P_1{*}d$ and $Y_2 = Y_0 - P_2{*}d$. As above, I also assume a Cobb-Douglas utility function $U(X, Y; \alpha, \beta) = X^{\alpha}Y^{\beta}$. Note that the slope (absolute value) of the tangent line of the indifference curve calculated at (X, Y) is $\alpha Y/(\beta X)$, and that the slope (absolute value) for the tangent line of the indifference curve at the kink (X_0, Y_0) is between P_1 and P_2. For the simplest case where $\alpha = \beta$,

$$U_1 < U_2 \Leftrightarrow X_1Y_1 < X_2Y_2$$

$$\Leftrightarrow \ (X_0 - d)*(Y_0 + P_1*d) < (X_0 + d)*(Y_0 - P_2*d)$$

$$\Leftrightarrow \ Y_0/X_0 > (P_1 + P_2)/2 + (P_2 - P_1)*d/(2X_0) \tag{14}$$

Note that since d is small relative to X_0, the second term in right hand of equation (14) is close to 0.

For the general case where $\alpha \neq \beta$, take the first order Taylor expansion for U_1 and U_2 around X_0.

$$U_1 = X_1^{\alpha}Y_1^{\beta} \approx X_0^{\alpha}Y_0^{\beta} - \alpha X_0^{\alpha-1}Y_0^{\beta}d + \beta X_0^{\alpha}Y_0^{\beta-1}P_1*d$$

$$U_2 = X_1^{\alpha}Y_1^{\beta} \approx X_0^{\alpha}Y_0^{\beta} + \alpha X_0^{\alpha-1}Y_0^{\beta}d - \beta X_0^{\alpha}Y_0^{\beta-1}P_2*d \tag{15}$$

Then $U_1 < U_2 \Leftrightarrow \alpha Y_0/(\beta X_0) > (P_1 + P_2)/2$ (16)

Note that $\alpha Y_0/(\beta X_0)$ is the slope (absolute value) of the tangent line of the indifference curve at the kink. When the slope of the tangent line is above $(P_1 + P_2)/2$ or close to P_2 (the slope of lower segment), consumers will be more likely to err up when X_0 is not available, which means that consumer demand is skewed to the left.

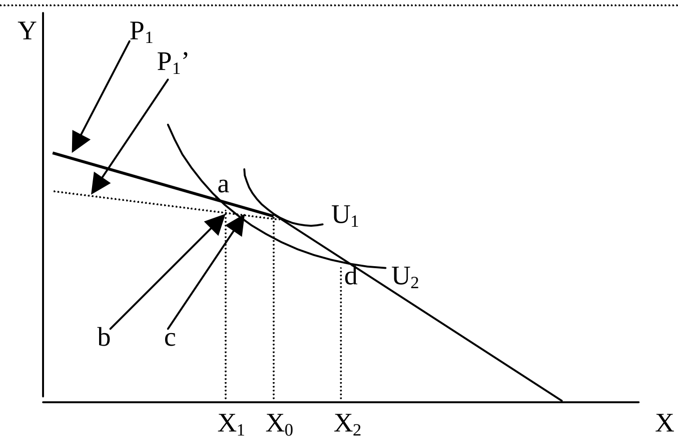


Figure 4. Piecewise-Linear Convex Budget Set

If P_1 decreases or P_2 increases, the kink becomes sharper, but the effects on skewness are different in these two cases: decrease in P_1 makes $\alpha Y_0/(\beta X_0)$ closer to P_2, so consumer demand is more likely skewed to the left, while increase in P_2 has an opposite effect. Consider a piecewise-linear convex budget set with two segments, as shown in Figure 4. The up segment moves down from P_1 to P_1'. Note that X_1 and X_2 here are different from above because they are not two equaldistance points to X_0 but two points (represented by a and d) on the same indifference curve U_2 under the old piecewise-linear budget constraint. From Figure 4 we know U(a) = U(c) = U(d) > U(b). Consumers, who maximize utility at the kink, are indifferent to X_1 and X_2 before the shift of the up segment. But after the shift, consumers will prefers X_2. Therefore, consumer demand is more likely to be skewed to the left if the kink is sharpened through lowering the up segment. This argument also holds for consumers located near the kink.[16]

[16] In the Monte Carlo study, the kink is sharpened through lowering the up segment. Tax Reform Act of 1986 (TRA86) also sharps the kink by lowering the first segment of the budget constraint faced by homeowners under 55.

Chapter 3. Monte Carlo Experiments with Simulated Data

3.1. Data Simulation

For the Monte Carlo experiments, I use the data and parameter values that are consistent with a two-error demand model for owner-occupied housing. The dependent variable is the log form of housing stock ($x = \log X$). Consumer's income (M_2) and rental cost for unit housing service (P_2) without capital gains tax are assumed to follow log-normal distribution: $\log(M_2)$ has mean 10.04 and standard deviation 0.35 (see appendix in Gabriel and Rosenthal (1991)), and P_2 mean 0.735 and standard deviation 0.062 (see Hoyt and Rosenthal (1992) Table 1). By definition, when the consumer buys down, $M_1 = M_2 - T P_2 X^*$ and $P_1 = P_2 (1 - T)$, in which X^* is the kink value, and T is the tax rate due to capital gains tax.[17] Although X^* depends on sale price of previous house, and T depends on the marginal income tax rate, I assume that $x^* = \log X^*$ follow a standard normal distribution and T to be a constant for simplicity.[18] In case measurement error (M_ERR) is skewed, I assume it follows a reversed gamma distribution with location parameter α and scale parameter λ.[19] The price elasticity a_1 is set to -1.32, the income elasticity a_2 is set to 0.07 (see Hoyt and Rosenthal (1990) Table A-1), and the constant term a_0 is set to 0 for simplicity.

1000 Monte Carlo repetitions were carried out and both OLS (use P_2 and M_2 as if there were no capital gains tax) and ML (use Quasi-Newton Optimization Method) estimators are examined. Note that in ML estimation I assume bivariate normal distributed error terms even if the measurement error is truly skewed. In each repetition, sample size (N) is set to 100, 500, and 1000, skewness to 0, -1, and –4, kink angle to 0, 0.25, and 1, and standard deviation for measurement error (M_SE) to 0.1, 0.2, 0.4.[20] Since I only require the overall log likelihood

---

[17] Kink angle $= P_2 / P_1 - 1 = T / (1 - T)$. In Table 1 of Hoyt and Rosenthal (1992), kink angle is defined as $R / (R - k * t) - 1$. In terms of my definition, $R = P_2$ and $R - k * t = P_1$, where t is the marginal income tax rate, and k is the capital gain tax rate (equal to 0.4 before 1986 and 1.0 after TRA86). The average kink angle before TRA86 in the data is 0.249 (I use 0.25 instead), so the average value of t is equal to 0.5. Substituting k with 1.0, we will have the average kink angle after TRA86 is 1.0.

[18] I choose three values for T: 0, 0.2, and 0.5. The corresponding kink angle are 0, 0.25, and 1.

[19] $M\_ERR = \alpha / \lambda - G(\alpha, \lambda)$ with mean = 0, variance $= \alpha / \lambda^2$, and skewness $= -2 / \text{sqrt}(\alpha)$. Note that when $\alpha > 1$ or $|\text{skewness}| < 2$, gamma distribution is bell shaped. The reason I use gamma distribution is that it is a built-in function in SAS. Other choices are truncated normal, Weibull, or Beta distributions.

[20] When skewness = 0, M_ERR follows normal distribution with mean 0. Graphs 1-2 give the histograms for

function to be positive for each observation, the ML procedure is unconstrained in terms of parameter restriction (see MaCurdy, Green, and Paarsch (1989) Table 2).

3.2. Monte Carlo Results

Tables 1-12 report summary statistics for OLS and ML estimates. Reported are the average estimates and the root mean squared errors. All the results are rounded to one-thousandth. Note that the root mean squared errors are not the standard errors for the average estimates. We can have standard errors for the average estimates by dividing root mean squared errors by square root of number of repetitions (=1000) which is around 32. Because the number of Monte Carlo repetitions is relatively large, the estimates should be fairly close to their theoretical values according to the Glavenko-Cantelli theorem (see Rao (1965)), so I can use the precision of the estimates (closeness to the true value) to investigate the factors' effects on the estimates. Since the value for income elasticity is set close to zero (=0.07), the summary statistics for a_2 isn't as informable as that for a_1. I will concentrate on the latter when I analyze the Monte Carlo results.

Tables 1-3 give the summary statistics for skewness = 0. First, let us take a look at the standard deviation of the estimates. Obviously, there is a trend in the standard deviation of the estimates. It increases from bottom to top and from left to right in each table for both ML and OLS estimates. This tells us a negative relationship between the dispersion of the estimates and sample size and a positive relationship between the dispersion and M_SE. Next, let us take a look at the estimates. In table 1 for kink angle = 0, OLS and ML estimates are the same, while in tables 2&3 for kink angle ≠ 0, OLS estimates are much smaller than ML estimates in magnitude, more different from the value set before the experiment than ML estimates, and biased. When kink angle increases from 0.25 (Table 2.) to 1.0 (Table 3.), the difference between OLS and ML estimates become larger, which tells us that the larger the kink angle is, the more biased the OLS estimates are (when kink angle = 0, the difference between MLE and OLS estimate is 0).

Tables 4-9 give the summary statistics for skewness ≠ 0. First, let's consider the standard deviation of estimates. In table 4&7, the standard deviation for estimates follows a same trend for N and M_SE as above, which suggests that skewness does not play a role in the dispersion of

---

skewness -1 and -4 with M_SE = 0.2. The histograms for M_SE = 0.1 or 0.4 are pretty same.

the. While in tables 4-5 and 8-9, the standard deviation for the estimate follows a trend for N as above, but it does not for M_SE. Next, let's consider the estimates. In tables 4 and 7, OLS and ML estimates are equivalent and BLUE in the presence of a linear budget constraint, although ML estimation isn't a proper procedure for skewness ≠ 0. While in tables 4-5 and 8-9, there isn't any clear pattern, which may be due to that too many factors play roles. In summary, the effect of sample size is fairly clear in any table: the larger N is, the less dispersed the MLE is. This suggests that I could fix N = 1000 and then discuss the effects of the other factors. Therefore, I produce Tables 10-12.

In tables 10-12, there is an obvious trend in the root mean squared errors: it increases from left to right in every table for both ML and OLS estimates. This trend suggests a positive relationship between the dispersion of estimates and M_SE. Note that this trend is not strict for nonzero skewness, which may be due to the model misspecification in ML estimation based on the bivariate normal distributed error terms.

Table 10 reports the results when kink angle = 0. In this table ML and OLS estimates are identical and pretty close to the value set before the experiment even when the measurement error follows a skewed distribution. This is consistent with the theoretical hypothesis that ML and OLS estimates are equivalent and both BLUE when the consumer faces a linear budget constraint.[21]

Tables 11-12 report the summary statistics for kink angle ≠ 0. In these two tables OLS gives biased estimate because consumers face kinked budget constraints. When skewness = 0, I find that ML estimates are pretty close to the value set before the experiment, and OLS estimates are much smaller than ML estimates in magnitude, different from the set value, and biased to 1, which is consistent with my expectation in section 1.4. When kink angle increases from 0.25 (Table 11) to 1.0 (Table 12) for skewness = 0, the difference between OLS and ML estimates (or the value set before the experiment) becomes larger. This tells us that the shaper the kink, the better ML estimation works than OLS estimation and the more biased OLS estimates.

Next, let's take a look at the estimates in Tables 11-12 when skewness ≠ 0. Because of the influence from skewness, most ML estimates are far different from the value set before the experiment and thus, inconsistent. For fixed value of skewness, the larger kink angle, the worse

[21] When I carefully check the computer results, I find that the difference between ML and OLS estimates takes place mostly on the ten thousandth when kink angle = 0.

ML estimates. This is consistent with my expectation about the interaction between kink angle and skewness in section 1.3. Also I find that ML estimates are not necessarily better than OLS estimates when kink angle (=1.0) is large and skewness is moderate (= -1.0).

A short summary of Monte Carlo studies is presented in the following. Both sample size and error variance have a significant effect on the dispersion of the estimates. ML and OLS estimates are equivalent and both BLUE when consumers face linear budget constraints. When there is a kink and the error is normally distributed, ML estimates are unbiased and consistent and far superior to OLS estimates. But when the error term is truly skewed in the presence of a kinked budget constraint, ML estimates, based on the assumption of normal error terms, are biased and inconsistent, and the larger kink angle, the more biased and inconsistent ML estimates.

Chapter 4. Estimation with Actual Data

4.1. Data and Variable Description

The primary data source used to estimate the two-error housing demand model was the 1981 SMSA file of the American Housing Survey (AHS).[22] The data included detailed information on the demographic and housing characteristics of over 75,000 families in fifteen SMSAs across the United States. From this group only families who moved in the twelve months prior to their interview were used. In addition, families had to be first time or previous homeowners that reported the purchase price of their home. These restrictions reduced the sample to 1306, of which 777 are first-time owners and 529 previously owned homes. The previous homeowners reported the purchase price of their home. Of the previous homeowners, roughly 85% bought up and 15% bought down.

Table 13 gives the descriptive statistics for all the variables. A price index for owner-occupied housing was used to deflate the sale prices of previous and current homes to determine a quantity measure. X and X* are the log forms of current and previous home values adjusted by the price index. The rental cost of housing services (P2) is defined as the cost to an owner-occupier of one dollar of housing services in the rental market, and M2 is the current consumer income. Both P2 and M2 are in log forms. P1 and M1 are the virtual price and income when consumers buy down.

Demographic variables in the model include previous tenure status (Pten), which equals 1 if the family is a previous owner and 0 if the family is a first time owner. If the household head is between 30 and 40 years old, Age3040 equals 1; if the household head is between 40 and 55 years old, Age4055 equals 1; if the household head is over 55, Age55 equals 1. The variables Dep1, Dep2, and Dep3, are also 0-1 dummies and equal 1 if there are 1, 2, and more than dependents, respectively, living at home, and 0 otherwise. Finally, the Race, Sex, and Marriage (Mar) variables are all 0-1 dummies and equal 1 if the household head is white, male, and married.

---

[22] Hoyt & Rosenthal (1990) used this data to estimate an owner-occupied housing problem. After 1983 AHS no longer reports the sale value of previous house, which is the kink value in the estimation. That's why I have to use the 1981 data instead of any recent AHS data. Because the data I used is a little different from theirs (even the OLS results are not identical), the estimation results are close, but slightly different from theirs too.

As suggested by Hoyt & Rosenthal (1990), housing price appreciation (π) is entered directly into the demand function as log ($1+\pi$). A probit model of whether families own or rent was also estimated prior to estimating the housing demand function for owner-occupied housing. A Mill's ratio term was then formed based on the estimated coefficients from the probit model, and included in the demand function as an additional regressor to control possible selection bias.

4.2. Estimation Results

Table 14 presents the estimation results when housing demand function was estimated under both linear and kinked budget constraints. Furthermore, under kinked budget constraint, measurement error is allowed to follow both normal and skewed distribution. The estimation was repeated separately for full sample, for previous homeowners, and for previous renters (recall that kinked budget constraint does not apply to previous owners over 55 years old and previous renters). Most of the coefficients are of the anticipated sign and significance. For instance, price elasticity of housing demand is negative while income elasticity is positive. Housing demand may increase with age and number of dependents, and white families, married couples and previous owners tend to purchase larger homes relative to black and single-headed families and previous renters.[23] In contrast, the sex of the household head does not appear to influence housing consumption, and the Mill's ratio term is insignificant. Finally, the housing price appreciation variable is positive and significant, which suggests that anticipated capital gains associated with housing price inflation has a positive impact on housing demand. Another interesting finding lies in the standard error of estimates. Since the degree freedom decreases by one as we switch from OLS model to Two-Error Normal model, and further by another one from Two-Error Normal model to Two-Error Normal & Gamma model, most estimates' standard errors increase in these three models.

As noted by Rosenthal, Duca, and Gabriel (1991), credit-constrained households exhibit different elasticities of demand for housing than less constrained households. In fact, credit rationing imposes two additional constraints on consumer's budget constraint: (1) an up limit on the share of current income owner-occupiers can spend on housing and (2) a minimum down-

[23] When (s)he is above 55 years old, the homeowner may demand less housing because his/her children may have a home themselves, and (s)he is also allowed a one-time exemption from payment of capital gains tax for the first $125,000 of capital gain.

payment ratio. A family is defined as credit rationed if either (1) or (2) is binding. When (1) is binding, the price elasticity of demand for housing is biased towards –1, while the income elasticity of demand is biased towards 1. When (2) is binding, the price elasticity is biased towards –1 while the income elasticity is biased toward 0. So the net effect of credit rationing on the estimated income elasticity is ambiguous, whereas the estimated price elasticity is unambiguously biased toward –1. Results from a variety of empirical studies suggest that, in the absence of credit constraints, the absolute value of the price elasticity is less than 1, which implies an upward bias to the price elasticity in case of credit rationing. Because the previous renters are, more likely, credit rationed, the price elasticity is biased upward to –1 comparing to previous owners who are less probably credit rationed.

Let us first take a look at the partial sample estimation results. Since previous renters face a linear budget constraint, OLS estimates for previous renters are BLUE no matter whether the error term is skewed or not. Since the price elasticity for previous renters is -0.72, we have a prior that the price elasticity for previous owners will be lower than 0.72 (in absolute value) because previous owners are less likely credit rationed. Previous owners under 55 years of age face piecewise-linear budget constraints, which leads OLS estimates for price and income elasticities biased to –1 and 0 respectively. Table 14 also reports the OLS estimates for previous renters and previous owners over 55 years old (recall that previous owners above 55 years old also face a linear budget constraint). Because in my data there are only 38 previous owners who are above 55 years old, the dummy variable terms for price and income are insignificant. Since two-error model recognizes the kinked budget constraint, ML estimates for two-error model will be unbiased and consistent if the assumption about the distribution of the error terms is correct. The estimate for price elasticity in two-error normal model is unbelievably large (= -1.64), which suggests that bivariate normality is not a reasonable assumption for the error terms. While in two-error normal & gamma model allowing skewness for the measurement error, the estimate for price elasticity (= -0.31) is consistent with our prior.

Table 14 also reports the estimation results for full sample. The estimated price elasticity (-0.59) in two-error normal & gamma model is consistent with our prior, while the estimated price elasticity (-1.31) in two error normal model is too large. Like the estimates in partial sample, all the estimates in two-error normal & gamma model are of the anticipated sign and significance, and especially the previous tenure variable is positive, which suggests that previous owners demand more housing than previous renters.

For both partial and full samples, σ_ε in two-error normal & gamma model is larger than that in two-error normal model, while σ_α in two-error normal & gamma model is smaller than that in two-error normal model. It seems that when measurement is truly skewed, two-error model assuming bivariate normality will underestimate the variance of measurement error while overestimating the variance of heterogeneity error.

Table 14 also reports the skewness of the measurement error and the percentage that the consumer errs on the up side. For previous owners, the skewness is –1.17 and 57.8% of the consumer will err on the up side. For the full sample, the skewness is –0.84 and 55.6% of the consumer will err on the up side. The histograms for such skewed reversed gamma distribution are given in Graphs 3 and 4. Note that since skewness of the measurement error is due to thinness in housing markets, the estimated skewness can be used as a measure for market thinness with known kink angle.

Although both Monte Carlo experiments and empirical estimation indicate that failing to allow for skewness in the demand for housing leads to biased estimates of the elasticities of demand when such skewness is actually present, the size of bias is quite different. In empirical estimation, controlling for skewness reduces the price elasticity of demand among previous owner-occupiers from 1.6 to 0.3, while in Monte Carlo experiments it only reduces from 1.32 to 1.20. The difference in size of bias may be due to the following reasons: (1) in empirical estimation both the error term and the independent variables are skewed, while in Monte Carlo experiments only measurement error follows a skewed distribution; (2) there are more explanatory variables in empirical estimation than in Monte Carlo experiments.

Conclusions

Previous empirical studies of consumer demand in the presence of piecewise-linear budget constraints typically ignore the possibility of error skewness in the estimation of two-error demand models. However, in the presence of thin markets, consumer demand is likely to be skewed when consumers face a kinked budget constraint. In this paper I derive necessary conditions for skewness to arise in the presence of thin markets and piecewise-linear budget constraints. I then conduct comprehensive Monte Carlo experiments for a two-error housing demand model to test the sensitivity of estimation results to skewness. I also estimate a two-error housing demand model using American Housing Survey (AHS) data allowing for both a kinked budget constraint and skewness, and compare results to those from a traditional kinked budget constraint model that assumes bivariate normality. Both Monte Carlo and estimation results based on the AHS data indicate that failing to allow for skewness can lead to substantially biased estimates of the elasticities of demand for housing when homeowners face a kinked budget constraint. In addition, 58 percent of previous owner-occupiers err on the up side when choosing their new home, while 42 percent err on the down side. Thus, housing demand is skewed.

References

Anas, Alex (1997). "Rent Control with Matching Economies: A Model of European Housing Market Regulation". Journal of Real Estate Finance and Economics, 15:1, 111-137 (1997).

Arnott, Richard J. (1988). "Housing Vacancies, Thin Markets, and Idiosyncratic Tastes". Queen's Institute for Economic Research Discussion Paper: 722 June 1988, pp. 40.

Aronsson, Thomas; Wikstrom, Magnus; Brannlund, Runar (1997). "Wage Determination under Non-linear Taxes: Estimation and an Application to Panel Data". Oxford Economic Papers; v49 n3 July 1997, pp. 404-18.

Aronsson, Thomas; Wikstrom, Magnus (1994). "Nonlinear Taxes in a Life-Cycle Consistent Model of Family Labor Supply". Empirical Economics; v19 n1 1994, pp. 1-17.

Barnett, Richard R.; Levaggi, Rosella; Smith, Peter (1992). "Local Authority Expenditure Decisions: A Maximum Likelihood Analysis of Budget Setting in the Face of Piecewise Linear Budget Constraints". Oxford Economic Papers; v44 n1 January 1992, pp. 113-34.

Billings, R.B., and Agthe, D.E.(1980). "Price Elasticities for Water: A Case of Increasing Block Rates". Land Economics, 56, 73-84.

Burtless, G., and Hausman, J. (1978). "The Effect of Taxation on Labor Supply: Evaluating the Gary Income Maintenance Experiment". Journal of Political Economy, 86, 1103-1130.

Caves, Richard E. (1978). "Industrial Organization and the Problem of Thin Markets". *Pricing Problems in the Food Industry (with Emphasis on Thin Markets).* Marvin L. Hayenga, editor. March 2-3, 1978.

Flood, Lennart; MaCurdy, Thomas (1992). "Work Disincentive Effects of Taxes: An Empirical Analysis of Swedish Men". Carnegie-Rochester Conference Series on Public Policy. December 1992.

Gabriel, Stuart; Rosenthal, Stuart (1991). "Credit Rationing, Race, and the Mortgage Market". Journal of Urban Economics 29, 371-379.

Glaeser, Edward L. and Kallal, Hedi D. (1997). "Thin Markets, Asymmetric Information, and Mortgage-Backed Securities". Journal of Financial Intermediation 6, 64-86 (1997).

Gray, Roger W. (1960). "The Characteristic Bias in Some Thin Futures Market". Food Research Institute Studies, Vol. 1, pp. 296-312, November, 1960.

Hausman, Jerry (1980). "The Effects of Wages, Taxes, and Fixed Costs on Women's Labor force Participation." Journal of Public Economics, 14 (October), 161-194.

Hausman, Jerry (1985). “The Econometrics of Nonlinear Budget Sets.” Econometrica 53 (Nov,1985), 1255-1282.

Hayenga, Marvin L.; Gardner, Bruce L.; Paul, Allen B.; and Houck, James P. (1978). “The Concept of a Thin Market”. *Pricing Problems in the Food Industry (with Emphasis on Thin Markets).* Marvin L. Hayenga, editor. March 2-3, 1978.

Hayenga, Marvin L., editor (1978). *Pricing Problems in the Food Industry (with Emphasis on Thin Markets).* March 2-3, 1978.

Heckman, J.J., and MaCurdy, T.E. (1981). “New Methods for Estimation Labor Supply Functions: A Survey”. Research in Labor Economics (Vol. 4), ed. R.G. Ehrenberg, Greenwich. CT: JAI Press.

Hoyt, William, and Rosenthal, Stuart (1990). "Capital Gains Taxation and the Demand for Owner-Occupied Housing." The Review of Economics and Statistics, Vol. LXXII, No. 1, February 1990.

Hoyt, William, and Rosenthal, Stuart (1992). "Owner-Occupied Housing, Capital Gains, and the Tax Reform Act of 1986." Journal of Urban Economics, 32 (1992), 119-139.

MaCurdy, Thomas (1992). "Work Disincentive Effects of Taxes: A Reexamination of Some Evidence." America Economics Review, Vol. 82 No. 2 (May 1992), 243-249.

MaCurdy, T., Green, D., and Paarsch, H. (1989). "Assessing Empirical Approaches for Analyzing Taxes and Labor Supply." Journal of Human Resources, 25(3): 415-90.

Martikainen, Teppo and Puttonen, Vesa (1994). “International Price Discovery in Finnish Stock Index Futures and Cash Markets”. Journal of Banking and Finance, v18, n5, October 1994, pp. 809-22.

Megdal, Sharon (1987). "The Econometrics of Piecewise-Linear Budget Constraints: A Monte Carlo Study." Journal of Business & Economics Statistics, April 1987, Vol. 5, No. 2.

Moffitt, Robert (1986). "The Econometrics of Piecewise-Linear Budget Constraints: A Survey and Exposition of the Maximum Likelihood Method." Journal of Business and Econimic Statistics 4 (July 1986), 317-328.

Nelson, Robert G. and Turner, Steven C. (1995). “Experimental Examination of a Thin Market: Price behavior in a Declining Terminal Market Revisited”. Journal of Agricultural and Applied Economics, 27(1), July, 1995: 149-160.

Pagano, Macro (1989). “Endogenous Market Thinness and Stock Price Volatility”. Review of Economic Studies, v56 n2 April 1989, pp. 269-87.

Pissarides, Christopher A. (1992). “Loss of Skill during Unemployment and the Persistence of

Employment Shocks". Quarterly Journal of Economics, v107, n4, November 1992, pp. 1371-91.

Powers, Mark J. (1978). "Thin Markets – A Regulatory Perspective". *Pricing Problems in the Food Industry (with Emphasis on Thin Markets).* Marvin L. Hayenga, editor. March 2-3, 1978.

Rao, C.R. (1965). *Linear Statistical Inference and Its Applications*. New York: John Wiley & Sons.

Rosenthal, Stuart S.; Duca, John V.; Gabriel, Stuart A. (1991). "Credit Rationing and the Demand for Owner-Occupied Housing." Journal of Urban Economics 30, 48-63 (1991).

Saleth, R. Maria; Braden, John B.; and Eheart, J. Wayland (1991). "Bargaining Rules for a Thin Spot Water Market". Land Economics, August 1991, 67(3): 326-39.

Sass, Tim R. (1987). " A Note on Optimal Price Cutting Behavior under Demand Uncertainty". The Review of Economics and Statistics, October 5, 1987.

Spanos, Aris (1986). *Statistical foundations of econometric modelling*. Cambridge University Press.

Taylor, L.D. (1975). "Demand for Electricity: A Survey". Bell Journal of Economics, 6, 74-110.

Taylor, L.D., Blattenbergre, G.R., and Rennhack, R.K. (1982). "Residential Energy Demand in the United States" (Vol. 1). Publication EA-1572, Palo Alto, CA: Electric Power Research Institute.

Terza, J.V., and Welch, W.P. (1982). "Estimating demand Under Block Rates: Electricity and Water". Land Economics, 58, 181-188.

Tomek, W.G. and Robinson, K.L. (1990). *Agricultural Product Prices*, 3rd ed. Ithaca, NY: Cornell Univ. Press.

Triest, Robert K. "A Monte Carlo Test of the Robustness of Alternative Estimators of Labor Supply Functions". Johns Hopkins Department of Political Economy Working Paper: 198 August 1987, pp. 32.

Wales, T.J., and Woodland, A.D. (1979). "Labour Supply and Progressive Taxes". Review of Economic Studies, 46, 83-95.

Appendix A. Mathematical Approximation to Compute Integral

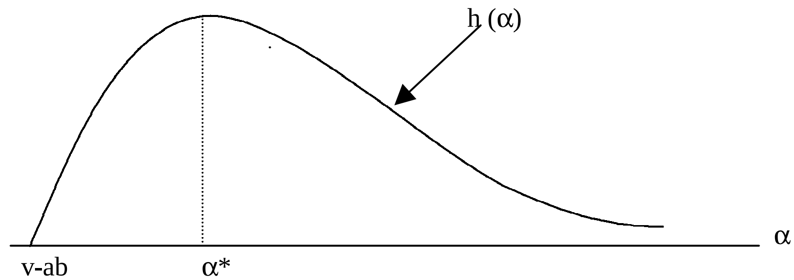


Figure 5. h (α)

From equation (7) we know that the bivariate distribution of v and α can be expressed as

$$f_{v,\alpha}(v, \alpha) = f_{u,\alpha}(ab + \alpha - v, \alpha) = g(ab + \alpha - v)\,\phi(\alpha/\sigma)/\sigma \qquad (17)$$

where $\alpha > (v - ab)$, g is the Gamma density function, ϕ is the standard normal density function, and a, b, and σ are three known parameters. Then for a given value of v, we can define a new function of α as

$$h(\alpha) = g(ab + \alpha - v)\,\phi(\alpha/\sigma)/\sigma, \text{ when } \alpha > (v - ab)$$

$$= 0, \text{ otherwise} \qquad (18)$$

Figure 5 gives the graph for h (α). Since the area under h (α) is finite [note the area under h (α) is equal to $f_v(v) = \int_{v-ab}^{\infty} [g(ab + \alpha - v)\,\phi(\alpha/\sigma)/\sigma]\,d\alpha$], h (α) will converge to zero when α approaches to infinity. Therefore, h (α) has one and only one maximum. Suppose h (α) maximizes at α*. If we take the first derivative of h (α) and let it equal to 0, then we have $\alpha^* = [m + \text{sqrt}(n)]/2$ in which $m = v - ab - \sigma^2/b$ and $n = m^2 - 4(1 - v/b)*\sigma^2$ [algebra is omitted here]. In equation (8) we need to compute

$$\text{Prob}_{1i} = \int_{v1-ab}^{\alpha 1} [g(ab + \alpha - v1)\,\phi(\alpha/\sigma)/\sigma]\,d\alpha$$

$$\text{Prob}_{2i} = \int_{\alpha 2}^{\infty} [g(ab + \alpha - v2)\,\phi(\alpha/\sigma)/\sigma]\,d\alpha$$

Prob_{1i} is equal to area under h (α) for $\alpha \in (v1 - ab, \alpha 1)$, and Prob_{2i} is equal to the area under h (α) for $\alpha > \alpha 2$. Therefore Prob_1 and Prob_2 can be approximately computed as:

$$\text{Prob}_{1i} = d^* \sum_{i=0}^{k} h(\alpha 1 - id)$$

$$\text{Prob}_{2i} = d^* \sum_{i=0}^{k} h(\alpha 2 + id)$$

where $\alpha 1 - kd > v1 - ab \geq \alpha 1 - (k + 1)*d$, $\alpha 2 + kd > \alpha^*$, $h(\alpha 2 + kd) < c$ and both d and c are small numbers (I chose d = 0.1 and c = 0.0000001 in my SAS program).

Note the smaller d and c, the more precise the estimation results, and the longer the computing time.

Appendix B. SAS Program for Monte Carlo Study

```
/* direct the output and log to new destination */
proc printto new print='c:\yan\output.txt' log='c:\yan\log.txt';
run;

option pagesize=60 ls=80 nodate pageno=1;

/* difine the library to output Monte Carlo results */
libname yan 'c:/yan';


%macro pwl(n,stderr,tax,a1,a2,mc,fname);

/*********************************************************

Macro "pwl" gives both OLS and ML estimation for a simulated housing demand dataset.

n--number of observations
stderr---standard deviation for measurement error
tax---capital gain tax rate on rent when buying down
a1---coefficient for rent
a2---coefficient for income
mc---number of Mante Carlo repetition

I set const=0 and heterogeneity error follow Normal(0, 0.09)for simplicity.

*********************************************************/

/* Data Generation */

/* Generate data for Income and Rent */
data inc_r;
  do i=1 to &n;
    /* log income for buying up */
    inc2=10.04+0.35*rannor(4157624);
    /* rent for buying up */
    rent2=log(0.735)+(0.062/0.735)*rannor(86580946);
    hstar=ranuni(88654867); /* kink value */
    /* income for buying down */
    inc1=log(exp(inc2)-&tax*exp(rent2)*exp(hstar));
    rent1=rent2+log(1-&tax); /* rent for buying down */
    output;
  end;
run;
```

```
%do j=1 %to &mc; /* start macro loop */

/* Generate data for error */
data error;
  do i=1 to &n;
    h_err=0.3*rannor(0); /* heterogeneity error */
    m_err=&stderr*rannor(0); /* measurement error */
    output;
  end;
run;

/* set data for house demand */
data house;
  merge inc_r error;
  by i;
  if &a1*rent2+&a2*inc2+h_err>hstar /* buy up */
    then house=&a1*rent2+&a2*inc2+h_err+m_err;
  else if &a1*rent1+&a2*inc1+h_err<hstar /* buy down */
    then house=&a1*rent1+&a2*inc1+h_err+m_err;
  else
    house=hstar+m_err; /* stay at kink */
run;


/* OLS Estimation */
proc reg data=house noprint outest=ols1;
  model house=rent2 inc2;
run;


/* Maximum Likelihood Estimation */

proc iml; /* start IML */

/* read SAS dataset into matrix */
  use house;
  read all var{house rent1 inc1 rent2 inc2 hstar} into
    data;

/* define variables */
  house=data(|,1|);
  rent1=data(|,2|);
  inc1=data(|,3|);
  rent2=data(|,4|);
  inc2=data(|,5|);
```

```
  hstar=data(|,6|);

/* define initial value for the parameter */
    /* incpt, a1, a2, seh, sem */
  initpar={0 &a1 &a2 0.3 &stderr};

/* define objective function */
  start F_PWL(par)
      global(house, rent1,inc1,rent2,inc2,hstar);

/* define parameters */
    a0=par[1];
    a1=par[2];
    a2=par[3];
    sigmah=par[4];
    sigmam=par[5];
    sigmav=sqrt(sigmah**2+sigmam**2);
    rho=sigmah/sigmav;
    prob=j(&n,1,0);
    i=1;

/* caculate probability for each observation */
    do while(i<=&n);

/* refer to Moffitt (1986) about h1,h2,z1,z2,s,t1,t2,r1,r2 */
      h1=a0+a1*rent1[i]+a2*inc1[i];
      h2=a0+a1*rent2[i]+a2*inc2[i];
      z1=(house[i]-h1)/sigmav;
      z2=(house[i]-h2)/sigmav;
      s=(house[i]-hstar[i])/sigmam;
      t1=(hstar[i]-h1)/sigmah;
      t2=(hstar[i]-h2)/sigmah;
      r1=(t1-rho*z1)/sqrt(1-rho**2);
      r2=(t2-rho*z2)/sqrt(1-rho**2);

      /* the probability to buy up */
      prob1=pdf('NORMAL',z1)*cdf('NORMAL',r1)/sigmav;
      /* the probability to buy down */
      prob2=pdf('NORMAL',z2)*(1-cdf('NORMAL',r2))
         /sigmav;
      /* the probability to locate on the kink */
      prob3=pdf('NORMAL',s)*(cdf('NORMAL',t2)
         -cdf('NORMAL',t1))/sigmam;

      prob[i]=prob1+prob2+prob3;
      i=i+1;
```

```
      end;

      if prob>0 then f=sum(log(prob));
      return(f);

   finish F_PWL;


   optn={1}; /* set to Maximize */

/* run quasi-newton procedure */
   call nlpqn(retcod,beta,"F_PWL",initpar,optn);

/* in case of nonconvergence, set both OLS and ML output = 0 */
   beta0={0 0 0 0 0};
   if retcod<0 then beta=beta0;
   edit ols1;
   intercep=0; rent2=0; inc2=0;
   if retcod<0 then replace;


/* create dataset to be used out of IML */
   cnames={"const" "rent" "inc" "seh" "sem"};
   create mldat1 from beta (|colname=cnames|);
      append from beta;

quit; /* quit IML */

/* formalize the Monte Carlo Result */

data olsdat; set ols1;
n=&n; skew=0; kink=&tax/(1-&tax); m_se=&stderr; mtd="OLS";
a0=intercep; a1=rent2; a2=inc2;
keep n skew kink m_se mtd a0 a1 a2;

data mldat; set mldat1;
n=&n; skew=0; kink=&tax/(1-&tax); m_se=&stderr; mtd="ML ";
a0=const; a1=rent; a2=inc;
keep n skew kink m_se mtd a0 a1 a2;
run;

/* output the results */
proc datasets nolist nodetails;
   append base=&fname data=olsdat;
   append base=&fname data=mldat;
run;
```

```
%end; /* end of macro loop */

%mend pwl; /* end of macro "pwl" */


%macro pwlsk(n,alpha,lamda,stderr,tax,a1,a2,mc,fname);

/*********************************************************

Macro "pwlsk" gives both OLS and ML estimation for a simulated housing demand dataset.

n--number of observations
(alpha,lamda)---measurement error which follows
            gamma(alpha,lamda)
stderr==sqrt(alpha)/lamda
a1---coefficient for rent
a2---coefficient for income
tax---capital gain tax rate on rent when buying down
mc---number of Mante Carlo repetition

I set const=0 and heterogeneity error follow Normal(0, 0.09)for simplicity.

note: skewness of mesurement error = 2/sqrt(alpha) and
     alpha>1

*********************************************************/

/* Data Generation */

/* Generate data for Income and Rent */
data inc_r;
  do i=1 to &n;
    /*income ($1000) for buying up */
    inc2=10.04+0.35*rannor(4157624);
    /* rent for buying up*
    rent2=log(0.735)+(0.062/0.735)*rannor(86580946);
    hstar=ranuni(88654867); /* kink value */
    /* income for buying down */
    inc1=log(exp(inc2)-&tax*exp(rent2)*exp(hstar));
    rent1=rent2+log(1-&tax); /* rent for buying down */
    output;
  end;
run;

%do j=1 %to &mc; /* start macro loop */
```

```
/* Generate data for error */
data error;
  do i=1 to &n;
    h_err=0.3*rannor(0); /* heterogeneity error*/
    /* measurement error */
    m_err=&alpha/&lamda-rangam(0,&alpha/&lamda)/&lamda;
    output;
  end;
run;

/* set data for house demand */
data house;
  merge inc_r error;
  by i;
  if &a1*rent2+&a2*inc2+h_err>hstar /* buy up */
    then house=&a1*rent2+&a2*inc2+h_err+m_err;
  else if &a1*rent1+&a2*inc1+h_err<hstar /* buy down */
    then house=&a1*rent1+&a2*inc1+h_err+m_err;
  else
    house=hstar+m_err; /* stay at kink */
run;


/* OLS Estimation */
proc reg data=house noprint outest=ols1;
  model house=rent2 inc2;
run;


/* Maximum Likelihood Estimation */

proc iml; /* stat IML */

/* read SAS dataset into matrix */
  use house;
  read all var{house rent1 inc1 rent2 inc2 hstar} into
    data;

/* define the variables */
  house=data(|,1|);
  rent1=data(|,2|);
  inc1=data(|,3|);
  rent2=data(|,4|);
  inc2=data(|,5|);
  hstar=data(|,6|);
```

```
/* define the initial values for the parameters */
    /* incpt, a1, a2, seh, sem;
  initpar={0 &a1 &a2 0.3 &stderr};

/* define the objective fuction */
  start F_PWL(par)
      global(house, rent1,inc1,rent2,inc2,hstar);

/* define the parameters */
    a0=par[1];
    a1=par[2];
    a2=par[3];
    sigmah=par[4];
    sigmam=par[5];
    sigmav=sqrt(sigmah**2+sigmam**2);
    rho=sigmah/sigmav;
    prob=j(&n,1,0);
    i=1;

/* caculate the probability for each observation */
    do while(i<=&n);

/* refer to Moffitt (1986) about h1,h2,z1,z2,s,t1,t2,r1,r2 */
      h1=a0+a1*rent1[i]+a2*inc1[i];
      h2=a0+a1*rent2[i]+a2*inc2[i];
      z1=(house[i]-h1)/sigmav;
      z2=(house[i]-h2)/sigmav;
      s=(house[i]-hstar[i])/sigmam;
      t1=(hstar[i]-h1)/sigmah;
      t2=(hstar[i]-h2)/sigmah;
      r1=(t1-rho*z1)/sqrt(1-rho**2);
      r2=(t2-rho*z2)/sqrt(1-rho**2);

      /* the probability to buy up */
      prob1=pdf('NORMAL',z1)*cdf('NORMAL',r1)/sigmav;
      /* the probability to buy down */
      prob2=pdf('NORMAL',z2)*(1-cdf('NORMAL',r2))
          /sigmav;
      /* the probability to locate on the kink */
      prob3=pdf('NORMAL',s)*(cdf('NORMAL',t2)
          -cdf('NORMAL',t1))/sigmam;

      prob[i]=prob1+prob2+prob3;
      i=i+1;
    end;
```

```
    if prob>0 then f=sum(log(prob));
    return(f);

  finish F_PWL;


  optn={1}; /* set to Maximize */

/* run quasi-newton procedure */
  call nlpqn(retcod,beta,"F_PWL",initpar,optn);

/* in case of nonconvergence, set both OLS and ML output = 0 */
  beta0={0 0 0 0 0};
  if retcod<0 then beta=beta0;
  edit ols1;
  intercep=0; rent2=0; inc2=0;
  if retcod<0 then replace;

/* create the data to be used out of IML procedure */
  cnames={"const" "rent" "inc" "seh" "sem"};
  create mldat1 from beta (|colname=cnames|);
    append from beta;

quit; /* quit IML */

/* formalize the Monte Carlo Results */

data olsdat; set ols1;
n=&n; skew=2/sqrt(&alpha); kink=&tax/(1-&tax); m_se=&stderr; mtd="OLS";
a0=intercep; a1=rent2; a2=inc2;
keep n skew kink m_se mtd a0 a1 a2;

data mldat; set mldat1;
n=&n; skew=2/sqrt(&alpha); kink=&tax/(1-&tax); m_se=&stderr; mtd="ML ";
a0=const; a1=rent; a2=inc;
keep n skew kink m_se mtd a0 a1 a2;
run;

/* output the results */
proc datasets nolist nodetails;
  append base=&fname data=olsdat;
  append base=&fname data=mldat;
run;

%end; /* end of macro loop */
```

```
%mend pwlsk; /* end of macro "pwlsk" */


%macro exe(n,mc,fname);

/*********************************************************

n = sample size
mc = number of Monte Carlo Repetations
fname = output file name

*********************************************************/

%pwl(&n,0.1,0,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,a2,   mc */

%pwl(&n,0.1,0.2,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */

%pwl(&n,0.1,0.5,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */

%pwl(&n,0.2,0,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,a2,   mc */

%pwl(&n,0.2,0.2,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */

%pwl(&n,0.2,0.5,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */

%pwl(&n,0.4,0,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,a2,   mc */

%pwl(&n,0.4,0.2,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */

%pwl(&n,0.4,0.5,-1.32,0.07,&mc,&fname)
/*  n,stderr,tax,a1,  a2,   mc */


%pwlsk(&n,4,20,0.1,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,20,0.1,0.2,-1.32,0.07,&mc,&fname)
```

```
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,20,0.1,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,10,0.2,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,10,0.2,0.2,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,10,0.2,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,5,0.4,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,5,0.4,0.2,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,4,5,0.4,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */



%pwlsk(&n,0.25,5,0.1,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,5,0.1,0.2,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,5,0.1,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,2.5,0.2,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,2.5,0.2,0.2,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,2.5,0.2,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,1.25,0.4,0,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */
```

```
%pwlsk(&n,0.25,1.25,0.4,0.2,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%pwlsk(&n,0.25,1.25,0.4,0.5,-1.32,0.07,&mc,&fname)
/*  n,alpha,lamda,stderr,tax,a1,a2,mc */

%mend exe; /* end of macro */



%exe(1000,100,yan.new) /* run the Macro */
    /* n, mc */


/* redirect output and log to default destination */
proc printto;
run;

quit;
```

Appendix C. SAS Program for Estimation of Two-Error Normal Model

```
option pagesize=55 ls=80 nodate pageno=1;

/*****************************************************

This program conducts a Maximum Likelihood Estimation
for a two-error demand model when the budget constraint
is piecewise-linear and the budget set is convex.

Dataset house1 has 19 variables, in which

house--response varible
hstar--kink value
p1--price for segment 1
y1--virtual income for segment 1
p2--price for segment 2
y2--virtual income for segment 2
k--dummy variable: 1 if the consumer faces kinked budget
   constraint; 0 if linear budget constraint
pten,agea,ageb,agec,dep1,dep2,dep3,race,sex,mar,pie,mill
 --social demographic variables

Two more coefficients are estimated by the program:

sigmah--standard error for heterogeinity error
sigmam--standard error for measurement error

********************************************************/

/* define the library which has the data */
libname yan 'c:/yan';

/* modify the data to be used in IML. you can
ignore this part if your data is ready */
data house1;
set yan.house;
if y1=. then delete;
if pten=1 and agec=0 then k=1;
else k=0;
if k=1 then hstar=log((pv1+pv2)/2);
else hstar=0;
drop obs pxj q taxrate cpi proptax rate
    pv1 pv2 s;
run;
```

```
proc iml; /* start IML */

/* read sas dataset into matrix */
   use house1;
   read all into data;
   nobs=nrow(data);

/* define the variables */
   agea=data(|,1|); ageb=data(|,2|);
   agec=data(|,3|); dep1=data(|,4|);
   dep2=data(|,5|); pie=data(|,6|);
   dep3=data(|,7|); race=data(|,8|);
   sex=data(|,9|); mar=data(|,10|);
   mill=data(|,11|); pten=data(|,12|);
   p2=data(|,13|); y2=data(|,14|);
   p1=data(|,15|); y1=data(|,16|);
   house=data(|,17|); k=data(|,18|);
   hstar=data(|,19|);

/* difine the objective function */
   start F_PWL(par)
         global(agea,ageb,agec,dep1,dep2,pie,
              dep3,race,sex,mar,mill,pten,
              p2,y2,p1,y1,house,k,hstar,nobs);

/* define the papameters */
      a0=par[1]; a1=par[2]; a2=par[3]; a3=par[4];
      a4=par[5]; a5=par[6]; a6=par[7]; a7=par[8];
      a8=par[9]; a9=par[10]; a10=par[11];
      a11=par[12]; a12=par[13]; a13=par[14];
      a14=par[15]; sigmah=par[16]; sigmam=par[17];
      sigmav=sqrt(sigmah**2+sigmam**2);
      rho=sigmah/sigmav;
      prob=j(nobs,1,0);
      i=1;

/* caculate the probability for each observation */
      do while(i<=nobs);

/* refer to Moffitt (1986) about h1,h2,z1,z2,s,t1,t2,r1,r2 */
         h2=a0+a1*p2[i]+a2*y2[i]+a3*pten[i]
           +a4*agea[i]+a5*ageb[i]+a6*agec[i]
           +a7*dep1[i]+a8*dep2[i]+a9*dep3[i]
           +a10*race[i]+a11*sex[i]+a12*mar[i]
```

```
       +a13*pie[i]+a14*mill[i];
      if k=0 then h1=h2;
      else
        h1=a0+a1*p1[i]+a2*y1[i]+a3*pten[i]
            +a4*agea[i]+a5*ageb[i]+a6*agec[i]
            +a7*dep1[i]+a8*dep2[i]+a9*dep3[i]
            +a10*race[i]+a11*sex[i]+a12*mar[i]
            +a13*pie[i]+a14*mill[i];
      z2=(house[i]-h2)/sigmav;
      z1=(house[i]-h1)/sigmav;
      t2=(hstar[i]-h2)/sigmah;
      t1=(hstar[i]-h1)/sigmah;
      s=(house[i]-hstar[i])/sigmam;
      r2=(t2-rho*z2)/sqrt(1-rho**2);
      r1=(t1-rho*z1)/sqrt(1-rho**2);

      /* the probability to buy up */
      prob1=pdf('NORMAL',z1)*cdf('NORMAL',r1)/sigmav;
      /* the probability to buy down */
      prob2=pdf('NORMAL',z2)*(1-cdf('NORMAL',r2))
          /sigmav;
      /* the probability to locate on the kink */
      prob3=pdf('NORMAL',s)*(cdf('NORMAL',t2)
          -cdf('NORMAL',t1))/sigmam;

      prob[i]=prob1+prob2+prob3;
      i=i+1;
     end;

    if prob>0 then f=sum(log(prob));
    return(f);

  finish F_PWL;

/* initial values for the parameters */
  initpar={9.56 -1.32 0.07 0.2 0.08 0.16 0.09
        0.05 0 0.11 -0.19 0.05 0.11 1.19
        0.04 0.33 0.16};

  optn={1 2}; /* set to Maximize */

/* run quasi-newton procedure */
  call nlpqn(retcod,beta,"F_PWL",initpar,optn);

/* run Finite Differences Approximation procedure to compute Hessian */
  call nlpfdd(f,grad,hessian,"F_PWL",beta);
```

```
/* define format of the output */
  rname={Const Price Inc Pten Age3040 Age4055 Age55 Dep1
       Dep2 Dep3 Race Sex Mar Pie Mills sigmah sigmam};
  cname={Initial_Value Estimate Stderr};

/* compute the standard error for each estimate */
  npar=ncol(rname);
  stderr=j(1,npar,0);
  cov=inv(-hessian);
  do i=1 to npar by 1;
    stderr[i]=sqrt(cov[i,i]);
  end;

/* output the results */
  variable=t(initpar//beta//stderr);
  title 'Normal Measurement Error, Sample Size=1306';
  print variable[rowname=rname colname=cname];
  print 'Value of Objective Function= ' f;

quit; /* quit IML */


quit;
```

Appendix D. SAS Program for Estimation of Two-Error Normal & Gamma Model

```
option pagesize=55 ls=80 nodate pageno=1;

/*****************************************************

This program conducts a Maximum Likelihood Estimation
for a two-error demand model when the budget constraint
is piecewise-linear and the budget set is convex.

Dataset house1 has 19 variables, in which

house--response varible
hstar--kink value
p1--price for segment 1
y1--virtual income for segment 1
p2--price for segment 2
y2--virtual income for segment 2
k--dummy variable: 1 if the consumer faces kinked budget
   constraint; 0 if linear budget constraint
pten,agea,ageb,agec,dep1,dep2,dep3,race,sex,mar,pie,mill
 --social demographic variables

Two more coefficients are estimated by the program:

sigma--standard error for heterogeinity error
alpha--location parameter for gamma distribution
beta--scale parameter for gamma distribution

*****************************************************/

/* define the library which has the data */
libname yan 'c:/yan';

/* modify the data to be used in IML. you can
ignore this part if your data is ready */
data house1;
set yan.house;
if y1=. then delete;
if pten=1 and agec=0 then k=1;
else k=0;
if k=1 then hstar=log((pv1+pv2)/2);
else hstar=0;
drop obs pxj q taxrate cpi proptax rate
```

```
    pv1 pv2 s;
run;


proc iml; /* statr IML */

/* read sas dataset into matrix */
   use house1;
   read all into data;
   nobs=nrow(data);

/* define the variables */
   agea=data(|,1|); ageb=data(|,2|);
   agec=data(|,3|); dep1=data(|,4|);
   dep2=data(|,5|); pie=data(|,6|);
   dep3=data(|,7|); race=data(|,8|);
   sex=data(|,9|); mar=data(|,10|);
   mill=data(|,11|); pten=data(|,12|);
   p2=data(|,13|); y2=data(|,14|);
   p1=data(|,15|); y1=data(|,16|);
   house=data(|,17|); k=data(|,18|);
   hstar=data(|,19|);

/* difine the objective function */
   start F_PWL(par)
         global(agea,ageb,agec,dep1,dep2,pie,
                dep3,race,sex,mar,mill,pten,
                p2,y2,p1,y1,house,k,hstar,nobs);

/* define the papameters */
      a0=par[1]; a1=par[2]; a2=par[3]; a3=par[4];
      a4=par[5]; a5=par[6]; a6=par[7]; a7=par[8];
      a8=par[9]; a9=par[10]; a10=par[11];
      a11=par[12]; a12=par[13]; a13=par[14];
      a14=par[15]; sigma=par[16]; alpha=par[17];
      lamda=par[18];
      prob=j(nobs,1,0);
      i=1;

/* caculate the probability for each observation */
      do while(i<=nobs);

/* refer to Moffitt (1986) about h1,h2,z1,z2,s,t1,t2,r1,r2 */
         h2=a0+a1*p2[i]+a2*y2[i]+a3*pten[i]
           +a4*agea[i]+a5*ageb[i]+a6*agec[i]
           +a7*dep1[i]+a8*dep2[i]+a9*dep3[i]
```

```
       +a10*race[i]+a11*sex[i]+a12*mar[i]
       +a13*pie[i]+a14*mill[i];
      if k=0 then h1=h2;
      else
         h1=a0+a1*p1[i]+a2*y1[i]+a3*pten[i]
           +a4*agea[i]+a5*ageb[i]+a6*agec[i]
           +a7*dep1[i]+a8*dep2[i]+a9*dep3[i]
           +a10*race[i]+a11*sex[i]+a12*mar[i]
           +a13*pie[i]+a14*mill[i];
      v2=house[i]-h2; v1=house[i]-h1;
      u2=hstar[i]-h2; u1=hstar[i]-h1;
      s=alpha*lamda-(house[i]-hstar[i]);

      /* the probability to buy up */
      prob1=INTEG1(alpha,lamda,sigma,v1,u1,0.1);
      /* the probability to buy down */
      prob2=INTEG2(alpha,lamda,sigma,v2,u2,0.1);
      /* the probability to locate on the kink */
      prob3=pdf('GAMMA',s,alpha,lamda)
          *(cdf('NORMAL',u2/sigma)
          -cdf('NORMAL',u1/sigma));

      prob[i]=prob1+prob2+prob3;
      i=i+1;
     end;

    if prob>0 then f=sum(log(prob));
    return(f);

  finish F_PWL;


  start INTEG1(alpha,lamda,sigma,v,ur,d);

/* This function approximately computes the integral of
gamma(alpha*lamda+u-v)*normal(u/sigma)/sigma on
(v-alpha*lamda)<u<ur. d is the interval. */

     sum=0;
     umin=v-alpha*lamda;

     do while (ur>umin);
       x=alpha*lamda+ur-v;
       temp=pdf('GAMMA',x,alpha,lamda)
           *pdf('NORMAL',ur/sigma)/sigma;
       sum=sum+temp;
```

```
        ur=ur-d;
      end;

      return(sum*d);

   finish INTEG1;


   start INTEG2(alpha,lamda,sigma,v,ul,d);

/* This function approximately computes the integral of
gamma(alpha*lamda+u-v)*normal(u/sigma)/sigma on u>ul.
d is the interval. */

      sum=0;
      umin=v-alpha*lamda;
      if ul<=umin then ul=umin+d;

      b=v-alpha*lamda-(sigma**2)/lamda;
      delta=b**2-4*(1-v/lamda)*(sigma**2);
      if delta<0 then umode=ul-1;
      else umode=(b+sqrt(delta))/2;

      do until (ul>umode & temp<0.00001);
        x=alpha*lamda+ul-v;
        temp=pdf('GAMMA',x,alpha,lamda)
           *pdf('NORMAL',ul/sigma)/sigma;
        sum=sum+temp;
        ul=ul+d;
      end;

      return(sum*d);

   finish INTEG2;


/* initial values for the parameters */
   initpar={9.56 -1.32 0.07 0.2 0.08 0.16 0.09
          0.05 0 0.11 -0.19 0.05 0.11 1.19
          0.04 0.33 1 0.16};

   optn={1 5}; /* set to Maximize */

/* set termination creteria to have more round than defult */
   tc={1000 6000};
```

```
/* set constraits for some parameters in case
they become negative or equal to 0 */
  con={. . . . . . . . . . . . . . . . 0.00001 0.00001
      0.00001,
      . . . . . . . . . . . . . . . . . .};

/* run Nelder-Mead Simplex procedure in order to increase
the speed to convergence */
  call nlpnms(retcod,beta,"F_PWL",initpar,optn,con,tc);

/* run Finite Differences Approximation procedure to compute Hessian */
  call nlpfdd(f,grad,hessian,"F_PWL",beta);

/* define format of the output */
  rname={Const Price Inc Pten Age3040 Age4055 Age55 Dep1
       Dep2 Dep3 Race Sex Mar Pie Mills Sigmah Alpha
       Lamda};
  cname={Initial_Value Estimate Stderr};

/* compute the standard error for each estimate */
  npar=ncol(rname);
  stderr=j(1,npar,0);
  cov=inv(-hessian);
  do i=1 to npar by 1;
    if cov[i,i]>0 then stderr[i]=sqrt(cov[i,i]);
  end;

/* output the results */
  variable=t(initpar//beta//stderr);
  title 'Gamma Measurement Error, Sample Size=1306';
  print variable[rowname=rname colname=cname];
  print 'Value of Objective Function = ' f;

quit; /* quit IML */


quit;
```

Table 1. SKEW= 0.0, KINK = 0.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.331 (0.379)	-1.331 (0.379)	0.075 (0.106)	0.075 (0.106)	-1.316 (0.434)	-1.316 (0.434)	0.069 (0.119)	0.069 (0.119)	-1.287 (0.560)	-1.287 (0.560)	0.068 (0.165)	0.068 (0.165)
500	-1.321 (0.167)	-1.321 (0.167)	0.071 (0.043)	0.071 (0.043)	-1.322 (0.180)	-1.322 (0.180)	0.070 (0.049)	0.070 (0.049)	-1.315 (0.270)	-1.315 (0.270)	0.066 (0.070)	0.066 (0.070)
1000	-1.320 (0.117)	-1.320 (0.117)	0.072 (0.30)	0.072 (0.30)	-1.324 (0.138)	-1.324 (0.138)	0.069 (0.034)	0.069 (0.034)	-1.314 (0.184)	-1.314 (0.184)	0.071 (0.049)	0.071 (0.049)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 2. SKEW= 0.0, KINK = 0.25
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.374 (0.378)	-1.182 (0.344)	0.074 (0.104)	0.060 (0.098)	-1.362 (0.442)	-1.183 (0.385)	0.070 (0.123)	0.056 (0.115)	-1.372 (0.723)	-1.187 (0.567)	0.083 (0.172)	0.067 (0.163)
500	-1.330 (0.176)	-1.217 (0.158)	0.071 (0.044)	0.070 (0.041)	-1.320 (0.192)	-1.215 (0.173)	0.069 (0.051)	0.068 (0.048)	-1.329 (0.304)	-1.215 (0.247)	0.070 (0.071)	0.069 (0.067)
1000	-1.327 (0.126)	-1.221 (0.110)	0.071 (0.029)	0.070 (0.028)	-1.329 (0.148)	-1.226 (0.127)	0.071 (0.034)	0.070 (0.032)	-1.329 (0.230)	-1.226 (0.186)	0.073 (0.050)	0.072 (0.047)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 3. SKEW= 0.0, KINK = 1.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.329 (0.382)	-1.127 (0.335)	0.073 (0.106)	0.057 (0.098)	-1.348 (0.437)	-1.139 (0.383)	0.074 (0.123)	0.057 (0.114)	-1.331 (0.625)	-1.085 (0.560)	0.078 (0.171)	0.057 (0.158)
500	-1.323 (0.179)	-1.195 (0.161)	0.068 (0.044)	0.067 (0.041)	-1.327 (0.209)	-1.200 (0.189)	0.067 (0.051)	0.066 (0.048)	-1.312 (0.291)	-1.186 (0.263)	0.074 (0.074)	0.072 (0.069)
1000	-1.320 (0.122)	-1.196 (0.109)	0.072 (0.030)	0.071 (0.028)	-1.315 (0.136)	-1.191 (0.122)	0.068 (0.035)	0.068 (0.032)	-1.318 (0.198)	-1.194 (0.177)	0.070 (0.051)	0.069 (0.047)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 4. SKEW = -1.0, KINK = 0.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.319 (0.368)	-1.319 (0.368)	0.068 (0.101)	0.068 (0.101)	-1.323 (0.348)	-1.323 (0.348)	0.069 (0.102)	0.069 (0.102)	-1.330 (0.415)	-1.330 (0.415)	0.074 (0.115)	0.074 (0.115)
500	-1.322 (0.159)	-1.322 (0.159)	0.069 (0.041)	0.069 (0.041)	-1.315 (0.160)	-1.315 (0.160)	0.070 (0.045)	0.070 (0.045)	-1.330 (0.186)	-1.330 (0.186)	0.067 (0.049)	0.067 (0.049)
1000	-1.321 (0.107)	-1.321 (0.107)	0.071 (0.028)	0.071 (0.028)	-1.316 (0.113)	-1.316 (0.113)	0.070 (0.029)	0.070 (0.029)	-1.319 (0.129)	-1.319 (0.129)	0.068 (0.033)	0.068 (0.033)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 5. SKEW = -1.0, KINK = 0.25
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.206 (0.330)	-1.147 (0.320)	0.063 (0.095)	0.059 (0.093)	-1.192 (0.328)	-1.179 (0.324)	0.063 (0.095)	0.062 (0.095)	-1.187 (0.384)	-1.182 (0.383)	0.061 (0.109)	0.060 (0.110)
500	-1.250 (0.150)	-1.219 (0.147)	0.065 (0.039)	0.065 (0.039)	-1.225 (0.157)	-1.220 (0.157)	0.069 (0.040)	0.069 (0.040)	-1.216 (0.176)	-1.215 (0.176)	0.068 (0.046)	0.068 (0.046)
1000	-1.251 (0.113)	-1.220 (0.110)	0.068 (0.027)	0.068 (0.027)	-1.221 (0.106)	-1.217 (0.106)	0.069 (0.027)	0.069 (0.027)	-1.229 (0.123)	-1.229 (0.123)	0.069 (0.032)	0.069 (0.032)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 6. SKEW = -1.0, KINK = 1.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.164 (0.338)	-1.113 (0.320)	0.063 (0.097)	0.059 (0.095)	-1.143 (0.321)	-1.137 (0.318)	0.048 (0.096)	0.047 (0.096)	-1.125 (0.369)	-1.123 (0.368)	0.055 (0.115)	0.055 (0.114)
500	-1.210 (0.148)	-1.183 (0.143)	0.069 (0.039)	0.069 (0.039)	-1.198 (0.148)	-1.196 (0.148)	0.068 (0.039)	0.068 (0.039)	-1.188 (0.171)	-1.188 (0.171)	0.068 (0.046)	0.068 (0.046)
1000	-1.225 (0.103)	-1.198 (0.100)	0.069 (0.027)	0.069 (0.027)	-1.200 (0.105)	-1.198 (0.105)	0.070 (0.027)	0.070 (0.027)	-1.201 (0.124)	-1.200 (0.124)	0.067 (0.031)	0.067 (0.031)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 7. SKEW = -4.0, KINK = 0.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.325 (0.352)	-1.325 (0.352)	0.073 (0.100)	0.073 (0.100)	-1.306 (0.381)	-1.306 (0.381)	0.064 (0.106)	0.064 (0.106)	-1.279 (0.549)	-1.279 (0.549)	0.068 (0.159)	0.068 (0.159)
500	-1.315 (0.163)	-1.315 (0.163)	0.069 (0.042)	0.069 (0.042)	-1.314 (0.172)	-1.314 (0.172)	0.066 (0.046)	0.066 (0.046)	-1.321 (0.248)	-1.321 (0.248)	0.069 (0.065)	0.069 (0.065)
1000	-1.307 (0.119)	-1.307 (0.119)	0.071 (0.029)	0.071 (0.029)	-1.320 (0.123)	-1.320 (0.123)	0.069 (0.031)	0.069 (0.031)	-1.314 (0.176)	-1.314 (0.176)	0.069 (0.045)	0.069 (0.045)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 8. SKEW = -4.0, KINK = 0.25
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.351 (0.338)	-1.172 (0.325)	0.069 (0.104)	0.057 (0.099)	-1.305 (0.385)	-1.175 (0.355)	0.068 (0.108)	0.058 (0.104)	-1.271 (0.612)	-1.122 (0.507)	0.064 (0.159)	0.053 (0.154)
500	-1.307 (0.149)	-1.220 (0.146)	0.067 (0.042)	0.066 (0.039)	-1.263 (0.179)	-1.221 (0.161)	0.070 (0.045)	0.069 (0.044)	-1.255 (0.266)	-1.209 (0.248)	0.067 (0.063)	0.067 (0.062)
1000	-1.301 (0.113)	-1.224 (0.110)	0.069 (0.029)	0.069 (0.028)	-1.260 (0.135)	-1.223 (0.118)	0.068 (0.030)	0.067 (0.029)	-1.258 (0.182)	-1.212 (0.163)	0.070 (0.043)	0.070 (0.043)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 9. SKEW = -4.0, KINK = 1.0
(Numbers in Parentheses Are Root Mean Square Errors)

N =	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
100	-1.281 (0.338)	-1.111 (0.309)	0.066 (0.102)	0.053 (0.097)	-1.262 (0.403)	-1.124 (0.349)	0.063 (0.106)	0.052 (0.100)	-1.256 (0.576)	-1.089 (0.507)	0.064 (0.150)	0.053 (0.144)
500	-1.273 (0.191)	-1.196 (0.151)	0.072 (0.043)	0.071 (0.041)	-1.266 (0.193)	-1.203 (0.162)	0.071 (0.043)	0.070 (0.042)	-1.240 (0.256)	-1.183 (0.239)	0.069 (0.062)	0.069 (0.061)
1000	-1.248 (0.147)	-1.197 (0.104)	0.071 (0.029)	0.069 (0.027)	-1.255 (0.131)	-1.191 (0.114)	0.071 (0.030)	0.070 (0.029)	-1.250 (0.180)	-1.200 (0.170)	0.068 (0.045)	0.068 (0.044)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 10. N = 1000, KINK = 0.0
(Numbers in Parentheses Are Root Mean Square Errors)

SKEW	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
0	-1.320 (0.117)	-1.320 (0.117)	0.072 (0.30)	0.072 (0.30)	-1.324 (0.138)	-1.324 (0.138)	0.069 (0.034)	0.069 (0.034)	-1.314 (0.184)	-1.314 (0.184)	0.071 (0.049)	0.071 (0.049)
-1.0	-1.321 (0.107)	-1.321 (0.107)	0.071 (0.028)	0.071 (0.028)	-1.316 (0.113)	-1.316 (0.113)	0.070 (0.029)	0.070 (0.029)	-1.319 (0.129)	-1.319 (0.129)	0.068 (0.033)	0.068 (0.033)
-4.0	-1.307 (0.119)	-1.307 (0.119)	0.071 (0.029)	0.071 (0.029)	-1.320 (0.123)	-1.320 (0.123)	0.069 (0.031)	0.069 (0.031)	-1.314 (0.176)	-1.314 (0.176)	0.069 (0.045)	0.069 (0.045)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 11. N = 1000, KINK = 0.25
(Numbers in Parentheses Are Root Mean Square Errors)

SKEW	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
0	-1.327 (0.126)	-1.221 (0.110)	0.071 (0.029)	0.070 (0.028)	-1.329 (0.148)	-1.226 (0.127)	0.071 (0.034)	0.070 (0.032)	-1.329 (0.230)	-1.226 (0.186)	0.073 (0.050)	0.072 (0.047)
-1.0	-1.251 (0.113)	-1.220 (0.110)	0.068 (0.027)	0.068 (0.027)	-1.221 (0.106)	-1.217 (0.106)	0.069 (0.027)	0.069 (0.027)	-1.229 (0.123)	-1.229 (0.123)	0.069 (0.032)	0.069 (0.032)
-4.0	-1.301 (0.113)	-1.224 (0.110)	0.069 (0.029)	0.069 (0.028)	-1.260 (0.135)	-1.223 (0.118)	0.068 (0.030)	0.067 (0.029)	-1.258 (0.182)	-1.212 (0.163)	0.070 (0.043)	0.070 (0.043)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

Table 12. N = 1000, KINK = 1.0
(Numbers in Parentheses Are Root Mean Square Errors)

SKEW	M_SE = 0.1				M_SE = 0.2				M_SE = 0.4			
	A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)		A1 (=-1.32)		A2 (=0.07)	
	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS	ML	OLS
0	-1.320 (0.122)	-1.196 (0.109)	0.072 (0.030)	0.071 (0.028)	-1.315 (0.136)	-1.191 (0.122)	0.068 (0.035)	0.068 (0.032)	-1.318 (0.198)	-1.194 (0.177)	0.070 (0.051)	0.069 (0.047)
-1.0	-1.225 (0.103)	-1.198 (0.100)	0.069 (0.027)	0.069 (0.027)	-1.200 (0.105)	-1.198 (0.105)	0.070 (0.027)	0.070 (0.027)	-1.201 (0.124)	-1.200 (0.124)	0.067 (0.031)	0.067 (0.031)
-4.0	-1.248 (0.147)	-1.197 (0.104)	0.071 (0.029)	0.069 (0.027)	-1.255 (0.131)	-1.191 (0.114)	0.071 (0.030)	0.070 (0.029)	-1.250 (0.180)	-1.200 (0.170)	0.068 (0.045)	0.068 (0.044)

Note: SKEW = Skewness, KINK = Kink Angle , N = Sample Size, M_SE = Standard Deviation for Measurement Error.

13. Data Description

Variable	N	Mean	Std Dev	Skewness	Kurtosis	Minimum	Maximum
X	1306	10.88047	0.449341	0.196856	4.368466	8.294	13.584
X*	529	10.88730	0.425468	-0.265407	2.256225	8.54617	12.3137
P2	1306	-0.26504	0.095087	-0.13786	-0.04357	-0.56667	0.02555
M2	1306	10.0413	0.466773	-1.46283	5.373468	6.6826	11.0568
P1	1306	-0.41887	0.162455	-0.34605	-0.00681	-1.0013	0.02555
Y1	1306	10.03399	0.465041	-1.47352	5.412975	6.6826	11.0423
PTEN	1306	0.405054	0.491091	0.387269	-1.85286	0	1
AGE3040	1306	0.640888	0.479924	-0.58803	-1.65676	0	1
AGE4055	1306	0.143185	0.350396	2.039761	2.163937	0	1
AGE55	1306	0.047473	0.21273	4.260992	16.18083	0	1
DEP1	1306	0.204441	0.403447	1.467416	0.153542	0	1
DEP2	1306	0.183767	0.387443	1.634915	0.673976	0	1
DEP3	1306	0.098775	0.298474	2.692635	5.258332	0	1
RACE	1306	0.093415	0.291125	2.797487	5.834864	0	1
SEX	1306	0.120214	0.325337	2.338305	3.472985	0	1
MAR	1306	0.790199	0.407322	-1.4271	0.036658	0	1
LOG(1+π)	1306	0.03012	0.025531	0.412354	-0.732	-0.01005	0.092579
MILL	1306	-0.55368	1.104499	0.127224	-1.56792	-3.5278	0.9999

Table 14. Estimation Results
(Numbers in Parentheses Are Standard Errors)

	Previous Renters OLS	Previous Owners OLS	Previous Renters & Owners Over 55	Previous Owners Two Error Normal	Previous Owners Two Error Normal & Gamma		Full Sample OLS	Full Sample Two Error Normal	Full Sample Two Error Normal & Gamma
Constant	8.3250	8.7866	8.3146	10.4752	9.3379		8.3779	9.1629	9.2889
	(0.5277)	(0.8144)	(0.5327)	(0.7388)	(0.7576)		(0.4346)	(0.4110)	(0.0004)
Price	-0.7198	-0.9582	-0.7127	-1.6380	-0.3127		-0.8700	-1.3062	-0.5943
	(0.2953)	(0.3450)	(0.2993)	(0.2507)	(0.2430)		(0.2099)	(0.1648)	(0.0628)
Price*Pten*Age55	---	---	0.1234	---	---		---	---	---
			(1.3879)						
Inc	0.2099	0.1792	0.2115	-0.0196	0.1557		0.1973	0.1047	0.1211
	(0.0571)	(0.0903)	(0.0578)	(0.0800)	(0.0801)		(0.0473)	(0.0433)	(4.5E-5)
Inc*Pten*Age55	---	---	0.0032	---	---		---	---	---
			(0.2658)						
Pten	---	---	0.2939	---	---		0.2590	0.2137	0.2400
			(2.3311)				(0.0290)	(0.0302)	(5.0E-5)
Age3040	0.0955	0.0183	0.0972	0.0003	-0.1952		0.0890	0.0803	0.0628
	(0.0351)	(0.0660)	(0.0357)	(0.0726)	(0.1321)		(0.0299)	(0.0305)	(4.3E-5)
Age4055	0.2219	0.0710	0.2237	0.1168	-0.0921		0.1703	0.1928	0.1929
	(0.0552)	(0.0700)	(0.0562)	(0.0766)	(0.1377)		(0.0388)	(0.0396)	(0.0372)
Age55	0.1150	0.0905	0.116	0.0109	-0.1291		0.1356	0.0933	0.0972
	(0.0826)	(0.0827)	(0.0841)	(0.0923)	(0.1402)		(0.0538)	(0.0559)	(0.0507)
Dep1	0.0155	0.0453	0.0084	0.0541	0.0162		0.0248	0.0303	0.0164
	(0.0379)	(0.0419)	(0.038)	(0.0458)	(0.0645)		(0.0280)	(0.0286)	(4.1E-5)
Dep2	-0.0095	0.0137	-0.0163	0.0292	0.0259		0.0019	0.0134	-0.0009
	(0.0450)	(0.0414)	(0.0455)	(0.0453)	(0.0653)		(0.0304)	(0.0310)	(0.0321)
Dep3	0.1359	0.0750	0.1388	0.1088	0.0534		0.1072	0.1217	0.1135
	(0.0535)	(0.0532)	(0.0539)	(0.0579)	(0.0810)		(0.0376)	(0.0384)	(0.0410)
Race	-0.2090	-0.0844	-0.2137	-0.1080	0.0974		-0.1800	-0.1948	-0.1646
	(0.0427)	(0.0679)	(0.0433)	(0.0746)	(0.1221)		(0.0355)	(0.0362)	(0.0391)

Table 14. Estimation Results (Continued)
(Numbers in Parentheses Are Standard Errors)

	Previous Renters OLS	Previous Owners OLS	Previous Renters & Owners Over 55	Previous Owners Two Error Normal	Previous Owners Two Error Normal & Gamma		Full Sample OLS	Full Sample Two Error Normal	Full Sample Two Error Normal & Gamma
Sex	0.0371 (0.0433)	0.0788 (0.0646)	0.0297 (0.0435)	0.0840 (0.0714)	0.2169 (0.1039)		0.0471 (0.0353)	0.0467 (0.0360)	0.0484 (0.0349)
Mar	0.0418 (0.0428)	0.1447 (0.0676)	0.0413 (0.0432)	0.2253 (0.0745)	0.2114 (0.0984)		0.0671 (0.0347)	0.1023 (0.0345)	0.0529 (4.2E-5)
Log(1+π)	0.0171 (0.0264)	1.6715 (0.6268)	1.2425 (0.5263)	1.9510 (0.6765)	1.5596 (0.9858)		1.3617 (0.4000)	1.5233 (0.4069)	1.4620 (0.4297)
Mills	0.0171 (0.0264)	-0.0246 (0.0287)	0.0158 (0.0264)	-0.0033 (0.0319)	0.0450 (0.0433)		0.0022 (0.0180)	0.0034 (0.0183)	0.0603 (4.3E-5)
σ	0.3750	0.3399	0.3818	---	---		0.3609	---	---
σ_α	---	---	---	0.3180 (0.0219)	0.2909 (0.0155)		---	0.3359 (0.0160)	0.3041 (0.0090)
σ_ε	---	---	---	0.1607 (0.0317)	0.3038		---	0.1443 (0.0319)	0.2453
Skewness	---	---	---	0	-1.1715		---	0	-0.8421
Percentage on the up side	---	---	---	50%	57.79%		---	50%	55.60%
Reversed Gamma(a)	---	---	---	---	2.9145 (0.0001)		---	---	5.6404 (0.0002)
Reversed Gamma(b)	---	---	---	---	0.1780 (3.7E-5)		---	---	0.1033 (3.8E-5)
Log Likelihood	---	---	---	-144.8912	133.4474		---	-486.6923	-219.4290
Sample Size	777	529	815	529	529		1306	1306	1306

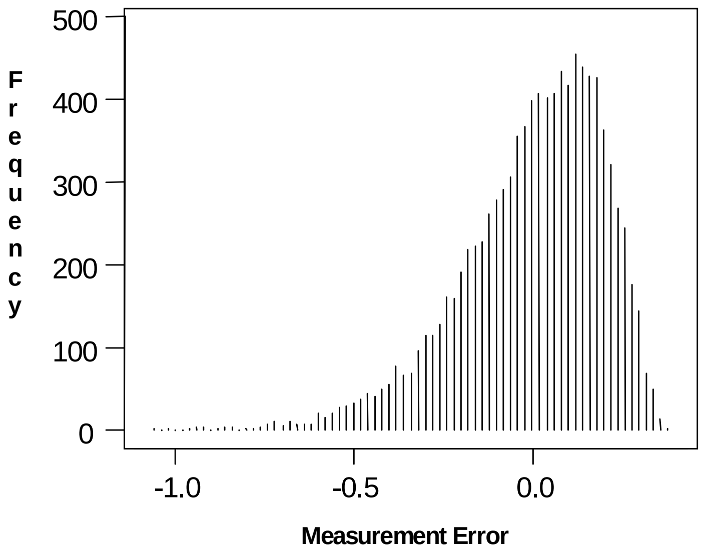
Table 15. Estimation Results from Hoyt & Rosenthal (1990)
(Numbers in Parentheses Are Standard Errors)

	Previous Renters OLS	Previous Owners OLS	Previous Owners Kink Range		Full Sample OLS	Full Sample Two Error Normal	Full Sample Kink Range
Constant	8.9376 (1.4419)	7.4333 (1.9743)	9.4118 (1.9305)		8.4734 (0.8439)	9.5588 (0.8685)	9.0823 (0.8496)
Price	-0.8544 (0.2443)	-1.0595 (0.3419)	-1.1793 (0.1670)		-0.9394 (0.1823)	-1.3198 (0.1520)	-1.0673 (0.1233)
Inc	0.1524 (0.1186)	0.2865 (0.1827)	0.0996 (0.1722)		0.1823 (0.0719)	0.0696 (0.0732)	0.1239 (0.0710)
Pten	---	---	---		0.2664 (0.0489)	0.2031 (0.0514)	0.2167 (0.0504)
Age3040	0.0857 (0.0556)	-0.0021 (0.0698)	0.0049 (0.0716)		0.0913 (0.0327)	0.0789 (0.0338)	0.0839 (0.0334)
Age4055	0.2138 (0.0619)	0.0843 (0.0585)	0.0673 (0.0629)		0.1765 (0.0377)	0.1620 (0.0392)	0.1696 (0.0391)
Age55	0.1146 (0.0850)	0.1469 (0.0722)	0.0309 (0.0767)		0.1378 (0.0502)	0.0909 (0.0513)	0.0943 (0.0509)
Dep1	0.0168 (0.0367)	0.0909 (0.0672)	0.0967 (0.0730)		0.0269 (0.0285)	0.0516 (0.0296)	0.0426 (0.0297)
Dep2	-0.0113 (0.0495)	0.0923 (0.0928)	0.0346 (0.0992)		0.0040 (0.0364)	0.0038 (0.0380)	0.0014 (0.0376)
Dep3	0.1388 (0.0513)	0.1490 (0.0967)	0.0733 (0.1015)		0.1089 (0.0392)	0.1085 (0.0409)	0.0975 (0.0396)
Race	-0.2159 (0.0398)	-0.1404 (0.0828)	-0.0867 (0.0888)		-0.1870 (0.0334)	-0.1925 (0.0349)	-0.1857 (0.0347)

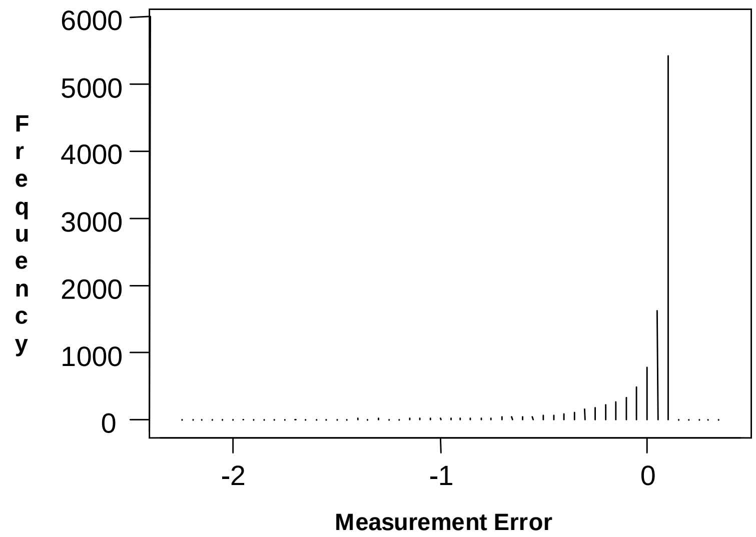
Table 15. Estimation Results from Hoyt & Rosenthal (1990) (Continued)
(Numbers in Parentheses Are Standard Errors)

	Previous Renters OLS	Previous Owners OLS	Previous Owners Kink Range		Full Sample OLS	Full Sample Two Error Normal	Full Sample Kink Range
Sex	0.0412 (0.0474)	0.0584 (0.0787)	0.0926 (0.0822)		0.0445 (0.0376)	0.0503 (0.0389)	0.0514 (0.0388)
Mar	0.0527 (0.0461)	0.1575 (0.0773)	0.1893 (0.0765)		0.0802 (0.0342)	0.1077 (0.0352)	0.0920 (0.0349)
Log(1+π)	1.0694 (0.5074)	1.4498 (0.6605)	1.3527 (0.7091)		1.0899 (0.3878)	1.1861 (0.4041)	1.0815 (0.3989)
Mills	0.0668 (0.1675)	-0.2506 (0.2360)	-0.0493 (0.2518)		-0.0114 (0.0956)	0.0380 (0.0997)	0.0285 (0.0983)
σ	0.3728	0.3359	0.3566		0.3598	---	0.3675
σ_α	---	---	---		---	0.3322 (0.0171)	---
σ_ε	---	---	---		---	0.1552 (0.0348)	---
Skewness	---	---	---		---	0	---
Percentile	---	---	---		---	50%	---
Reversed Gamma(a)	---	---	---		---	---	---
Reversed Gamma(b)	---	---	---		---	---	---
Log Likelihood	---	---	-235.8909		---	-490.9812	-579.632
Sample Size	777	529	529		1306	1306	1306

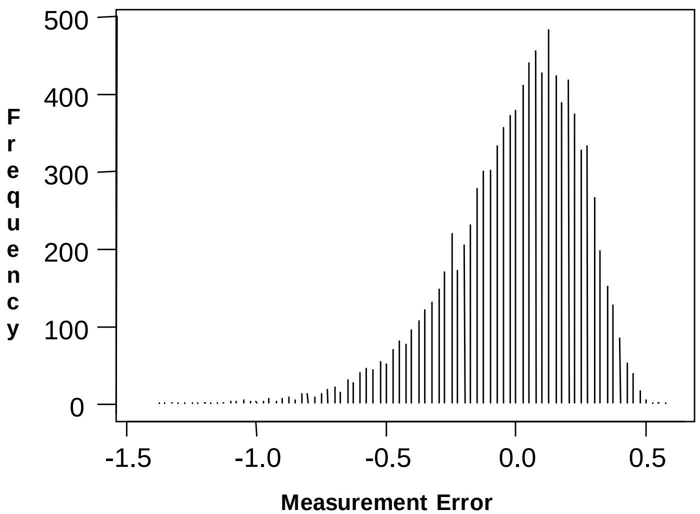
Graph 1. Stderr = 0.2, Skewness = -1.0



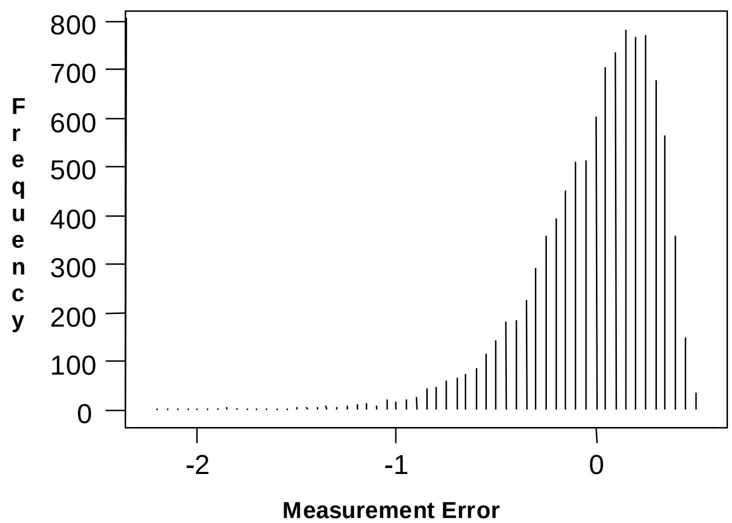
Graph 2. Stderr = 0.2, Skewness = -4.0



Graph 3. Stderr = 0.2453, Skewness = -0.8421



Graph 4. Stderr = 0.3038, Skewness = -1.1715



Yan, Zheng (John)

1214 University City Blvd, #H91, Blacksburg, VA 24060.
(415) 898-2255(office), (540) 552-1482(home) Email: zyan@vt.edu

OBJECTIVE: **Economist/Econometrician/Statistician**

EDUCATION: **Virginia Polytechnic Institute & State University (VPI&SU),** Blacksburg, VA

- **Ph.D., Economics**, expected July 1999
  Concentration: **Applied Econometrics**
  Dissertation Title: "Piecewise-Linear Budget Constraints, Thin Markets, and ErrorSkewness"
  Adviser: Dr. Stuart Rosenthal

- **MS, Statistics**, December 1998.
  Passed qualifying exam at Ph.D. level, May 1998

Nankai University, Tianjin, China
- MA, Economics, July 1995.
- BA, Economics, July 1992.

EXPERIENCE: **Paper Presentation in Eastern Economic Association Annual Meetings**, Boston, Mar 12, 1999.
Paper Title: Piecewise-Linear Budget Constraints, Thin Markets, and Error Skewness.

Dissertation: Piecewise-Linear Budget Constraints, Thin Markets, and Error Skewness,
Department of Economics, VPI&SU, February 1998 **-** present
- Literature review and theoretical motivation for skewness in the presence of kinked budget constraint and thin market
- Monte Carlo studies to test the sensitivity of the estimation results to skewness
- Estimation of kinked budget constraint problem allowing for skewness. Comparison of those results to the standard bivariate normal assumptions

A Dynamic Model of Quality of Life in a System of Cities, Department of Economics, VPI&SU, Summer 1998, Supervisor: Dr. Stuart Rosenthal
- Programmed with EVIEWS for an Error Correction Model to analyze quality of life
- Applied Error Correction Model to the annual data (1977-1991) for 39 cities from the American Housing Survey and the Current Population Survey

Model Selection: Women's Labor Supply Analysis. Department of Statistics, VPI&SU. Fall 1997.
Used various kinds of methods of analysis (Stepwise Selection, R-square, CP statistics, PRESS, Residual Analysis, Collinearity Diagnostics, and Detection of influential observations) to select an appropriate subset of variables that best explain women's labor supply.

A Multinomial Logit Household Model for Rural Surplus Labor Force in China, Department of Economics, VPI&SU, May 1997.
Used Multinomial Logit Model (MNL) to establish Households Decision-Making Model to analyze Chinese rural labor supply.

A Misspecification Test on a Simultaneous-Equations System, Department of Economics, VPI&SU. Fall 1996.
- Carried out a comprehensive misspecification test (including test for Normality, Functional Form, Homoskedasticity, Parameter Stability and Independence) on a two-equation system for commercial loans
- Suggested an appropriate respecification

Teacher Assistant, Tutor
Dept of Economics, VPI&SU, Supervisor: Dr. William Hyde. August 1995 - December 1998

Economist (Practical Training)
Junan Security Company, Shenzhen, China, September - December 1994

SKILLS: **Applied Economic and Statistical Modeling**

- Regression Methods
- Time Series and Forecasting Methods
- Limited Dependent and Qualitative Variables Methods
- Panel Data Methods
- Multivariate Methods

SAS Programming

- Data Management: BASE, MACRO, SQL, ACCESS, and FSP
- Statistical Analysis: STAT, IML, MACRO, and ETS
- Expertise in IML programming
- Programming in all platforms: PC, Unix, and Mainframe

Other Computer Skills

- Programming Languages: C, C++, Visual Basic, and SQL
- Statistical Software: EVIEWS, MINITAB, GAUSS, and JMP
- Operating Systems: Windows95&98, Unix (Sun), and MVS
- Application Software: MSWord, Excel, PowerPoint

Language Skills

- English: Strong communication and writing skills
- Chinese: Native language. Best presentation and writing skills

EMPLOYMENT: William E. Wecker Associates, Inc., 505 San Martin Dr., Novato, CA 94945.
Beginning on Aug 23, 1999.